Temporal Resolution of Uncertainty, the Investment Policy of Levered Firms and Corporate Debt Yields

Alexander S. Reisz[*]   Kose John[**]

JEL: G13, G31, L14

Last revised: March 6, 2002

[*] Assistant Professor, Finance and Economics Department, Zicklin School of Business, Baruch College (CUNY), 1 Bernard Baruch Way, Box B10-225, New York, NY 10010. Tel: (646) 312 3518, Fax: (646) 312 3451, E-mail: Alexander_Reisz@baruch.cuny.edu. This article is adapted from the first essay of my dissertation, written at the Stern School of Business, New York University.

[**] Corresponding author. Charles William Gerstenberg Professor of Banking and Finance, Stern School of Business, New York University, 44 W. 4th Street, Suite 9-190, New York, NY 10012. Tel: (212) 998 0337, Fax: (212) 995 4233, E-mail: kjohn@stern.nyu.edu.

1

Temporal Resolution of Uncertainty, the Investment Policy of Levered Firms and Corporate Debt Yields

Abstract

This paper attempts to link the agency literature (concerned with the fact that tensions between bondholders and shareholders may trigger suboptimal investment decisions) with the one dealing with temporal resolution of uncertainty (TRU). We consider here how the speed of resolution of the uncertainty characterizing the firm's operations affects the risk-shifting behavior of a shareholder-aligned manager. It is assumed that investors are risk neutral and that the return on the risky technology is normally distributed. It is shown that the speed of TRU affects monotonically the extent of risk shifting as well as bond yields, even after optimal contracts mitigating deviations from the first-best investment policy have been written. In particular, the optimal investment-restricting covenant is endogenously characterized. Empirical implications are derived and discussed.

1.   Introduction

A relatively recent review of the work done in the field of corporate finance over the last twenty five years ends with the conclusion that one of the biggest challenges facing theorists and empiricists alike is the valuation of "firms whose major asset consists of human rather than physical capital, prominent examples being in computer software and film production, […] for the paradigm example underlying most of our theoretical models is the manufacturing firm which dominated the growth of the economy around mid-century"[1]. More generally the technologies used by firms have evolved, as the industry groups represented in the S&P500, ever more heavily weighted towards pharmaceutical, banking, media and technology firms, testify. In particular, we believe that alongside human capital, the speed at which uncertainty is resolved for those firms/industries has become a more important factor of their investment policy and the yield demanded on their debt.

The relevance of the concept of temporal resolution of uncertainty (in the sequel: TRU) for corporate finance becomes intuitive as soon as one realizes that both sides of the balance sheet will depend heavily on it: the capital budgeting process will have to take into account whether more uncertainty affects short-term or long-term cash flows to figure out the right discount rates to use as well as the extent to which incoming information may change expectations about these cash flows. This will in turn determine the extent to which a firm can rely on debt as well as the type of debt it will use[2]. But the very nature of the firm's liabilities will affect its investment policy. This is certainly even more the case when we move from manufacturing firms to technology and human capital intensive firms, whose operations are not only arguably more risky, but also suffer from a more delayed resolution of uncertainty.

Indeed, a large strand of the corporate finance literature has focused on so-called "agency" problems: an agent/manager may undertake actions that are suboptimal from the principal's point of view. This may be due to a difference in their respective utilities or their information sets and affects as well the *level* as the *nature* of the investment. For instance, a shareholder-aligned manager may decide not to undertake a valuable investment project because proceeds have to be shared with bondholders (especially in the states of nature where the firm is barely solvent) and decide to strategically default instead (see Myers (1977)). Alternatively, the manager may play on the riskiness of the firm's investments. She may invest in assets that are too risky (or even in negative-NPV projects) to maximize the option value of equity (see for instance Jensen and Meckling (1976)) or, conversely, invest too conservatively. This would be the case if she is privately aware of growth options that would be lost in case of default (see Brito and John (2002)) or if she is not totally shareholder-aligned and is concerned about the loss of private benefits of control or reputation (especially if it is hard to tell whether a firm went bankrupt because of bad luck or bad management). However, the agency literature has focused on how

---

[1] Brennan (1995), p.18.

[2] See Goswami, Noe and Rebello (1995) for an attempt to design the type of debt a firm should use depending on whether more uncertainty surrounds short-term or long-term cash flows.

these agency games vary depending on concrete characteristics of the firm (bankruptcy costs, debt-related tax shields, level of tangible assets etc.) in order to obtain testable empirical implications and has left the effect of some more abstract concepts, such as temporal resolution of uncertainty, a yet… unresolved issue.

On the other hand, the literature dealing with temporal resolution of uncertainty looks at the firm as a black box producing cash flows. For instance, the pioneering work by Epstein and Turnbull (1980) showed that when the uncertainty about a cash flow is resolved *earlier* (in the sense that more information is available about it at an intermediate date), the return demanded by investors who have a CARA utility function in an environment where returns are multivariate normal is *larger* than the riskless rate: the expected release of information has rendered the holdings risky. However, Epstein and Turnbull's discussion is restricted to studying the effect of temporal resolution of uncertainty on equilibrium market prices and on the optimal production of information. In their model, after the manager has conducted the experiment that yields some early resolution of uncertainty, "the firm communicates the [result of the experiment] truthfully" and "no production decisions are taken after the experiment results become known" (p. 628). They are aware that there may be "a moral hazard in that actions by management may not be in the best interest of the owners and will decrease the market value of their holding" and that the firm may release spurious information, but consider that insisting upon monitoring of managerial activities will minimize the associated costs.

Nabar, Stapleton and Subrahmanyam (1988) took the issue into the realm of corporate finance by studying how the value of corporate debt (and hence its yield) is affected when the speed of uncertainty resolution varies. However, their model assumes a context in which "the assets of the firm are in place and future risky investment decisions have already been taken by the firm" (p. 224). As a consequence, Modigliani and Miller's (1958) Capital Structure Irrelevance Theorem still holds and "hence the leverage employed by the firm merely determines the split of total value between debt and equity and not the total value of the firm itself" (p. 225). Their model thus ignores agency problems and their effect on debt value and investment policy.

There have been a couple of articles looking at agency problems under the light of temporal resolution of uncertainty (see John (1987) and John and Ronen (1990)). They present a relatively rigid framework (two possible outcomes for the investment project and a manager's information set consisting of only the probability of the high outcome and/or its magnitude) and thus can only offer the conclusion that "the financing choices and the investment policy can be materially affected by the timing of financing and investment *vis-à-vis* crucial resolutions of uncertainty in the underlying technology"[3]. In particular, the relation between the speed of resolution of uncertainty and corporate bond yields is not studied. More generally, we don't believe any research has been done that allows the pattern of TRU to evolve along a continuous range and derives empirical implications on investment policy and corporate debt yields in an asymmetric information/agency games framework. We set out to fill this gap.

We offer a model where managers not only know the probability of success or the magnitude of the final cash flows, but observe at an intermediate date a private signal, which gives them partial but richer information about the return on the risky technology in which they can invest. Outsiders only observe a noisy version of this signal and rely on it to enforce debt covenants, trying to limit the distortions between the actual investment policy chosen by the shareholder-aligned manager and the one that would maximize firm value. We endogenously characterize the optimal covenants mitigating these suboptimal investment choices. This will lead us to various empirical implications, explaining, at least partially, why leverage and corporate bond yields differ consistently across industries.

The article is organized as follows: Section 2 posits the model and states the assumptions. Section 3 solves the problem, first when outsiders are not allowed to contract (or do not choose to), and then in the more general case when they can do so. We endogenously derive optimal covenants and when it is optimal to contract at all. Section 4 derives empirical implications and confronts them to the existing literature and Section 5 concludes. All proofs beyond the most trivial ones are relegated to the Appendix and numerical simulations and illustrating figures appear in the end.

2.   Model and Assumptions

We consider here a three-date (t=0,1 or 2), two-period model[4]. The sequence of events is as follows:

- At t=0, the entrepreneur, who owns the rights to a firm but does not have enough capital to finance it, sells claims consisting of debt and equity to outside investors. The debt is sold entirely to outsiders, while the entrepreneur may retain some of the equity or sell all of it[5]. In both cases, he has an incentive to maximize the combined value of both types of claims and therefore attaches to them firm-value-maximizing covenants.

- At t=1, the manager of a firm with cash resources of I is faced with two possible investments: a riskless one, yielding the (gross) riskfree rate r_2 (known at t=1) and a risky project that yields the (gross) stochastic rate θ. She makes her investment decision based on the observation of a private signal X_1. This signal is assumed to

---

[3] John (1987), p. 638.

[4] This discrete finite-period model is used not only to retain parsimony and elegance in a multi-period model, but also to have more flexibility in the correlation structure of cash flows: using a continuous time diffusion-type process would force us into a situation where we could not distinguish between the "pure" correlation element ρ and the time element $\sqrt{t/t'}$ that enters the covariance between two cash flows happening at different points in time t and t', t<t'.

[5] This insider contribution could be endogenized by trading off the reduction in contracting costs with the opportunity costs of supplying this capital, but we will abstain from it in this paper.

be bivariate normally distributed with the return available on the risky technology and hence gives some information to the manager about the probabilistic properties of the latter (we will denote ρ_m the correlation between X_1 and θ). Based on the observation of a particular realization x_1, the manager, a Bayesian decision maker, allocates her cash between the risky technology (in which she invests a fraction Q) and riskless Treasury bills in which she puts the remainder I-Q. At the time the manager observes the signal X_1, outsiders observe a noisy version of it, $Y_1 = X_1 + \varepsilon$, where $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ is white noise, independent of X_1 and θ; we will denote ρ_b the correlation between Y_1 and θ).

- At t=2, the realization of the risky technology θ is revealed to everybody and the firm is liquidated (proceeds are distributed according to the prewritten financial agreement).

The signal X_1 could be the result of a performance test on a prototype or a marketing survey; alternatively, it could be a t=1 cash flow, but we would have to impose the restriction that it is not contractible upon (or at least not verifiable by an outside court). We further assume that dividends are not allowed, in which case the shareholder-aligned manager would distribute as much as possible at t=1. Contracts can only be written (and enforced) based on the realization y_1 of the noisy signal Y_1, observable by everybody, not based on the realization x_1 of the more precise signal X_1, observed only by the manager. This asymmetry of information is at the heart of the agency problem and creates a problem of incomplete contracting in relation to the manager's investment choices. Our model can be summarized schematically as in Figure 1.

[Insert Figure 1 here]

"Temporal resolution of uncertainty" is defined in much the same way as in Epstein and Turnbull (1980): we say that the more (less) informative the experiment concerning the final cash flows is, the earlier (later) the prior uncertainty about cash flows is resolved. Epstein and Turnbull (1980) show that in the case of jointly normal distributions, the correlation coefficient is a sufficient statistic for the informativeness of the experiment: an experiment X is said to be more informative about θ than about θ' if corr(X,θ)$\geq$corr(X,θ'). This role of ρ is best seen if we consider the proportion of the uncertainty of θ that is resolved by observing X_1 or Y_1: $1 - \text{var}(\theta \mid X_1) / \text{var}(\theta) = \rho_m^2$ and $1 - \text{var}(\theta \mid Y_1) / \text{var}(\theta) = \rho_b^2$. Hence, we shall say that there is earlier resolution of uncertainty the larger (in absolute value) ρ, the correlation coefficient between the signal observed and θ.

In other words, if ρ is large, there is little more to be learnt at t=2 and most of the uncertainty has indeed been resolved in the first period. If ρ is low, most of the uncertainty is shifted towards the second period. We will consider only positive values of ρ, since we look at it as the correlation coefficient between θ and a random variable that is perfectly correlated to the signal and which tells you how informative this signal is about θ (for instance, a signal that has, say, a -0.9 correlation with θ would be extremely informative; in our model, this case would be summarized with ρ=0.9).

We need here to clarify a few assumptions of our model:

- *Assumption 1*: Investors are risk-neutral. This stems either from the absence of arbitrage (ensuring the existence of a risk-neutral measure), or from the combination of certain preferences and distribution properties for the underlying variables[6]. As a consequence, the firm's securities are priced on the basis of the relevant conditional expectation of the cash flows (with respect to the state price density function).

- *Assumption 2*: The firm has two types of marketed claims outstanding, debt and equity. The debt has the form of a pure discount bond of promised payment F which matures at t=2. Note that since the return on the investment is normally distributed, it has a positive probability of being arbitrarily negative and therefore any amount of outstanding debt is risky. Once the final cash flow is realized, the firm is obliged to pay debt claims, if possible. If the cash flow is insufficient to meet debt obligations, the firm goes bankrupt and its assets are turned over to bondholders. We will abstain from considering bankruptcy costs or the tax advantage of debt financing.

- *Assumption 3*: It is not possible for the manager to issue further debt at t=1 after observing the result of the experiment, nor is it allowed to invest a negative amount in the risky project, putting the proceeds in the riskfree asset ("short-selling" the risky project). We also assume that $E(\theta) > r_2$.

- *Assumption 4*: The information available to the manager at t=1 is restricted to a signal X_1[7].

- *Assumption 5*: The manager acts to maximize the wealth of current shareholders (i.e., under asymmetric information, the true value of their claims conditional on the private information). In a rational expectations equilibrium, debtholders as well as stockholders will correctly anticipate, at t=0, the effect of debt structure and temporal resolution of uncertainty on the chosen investment strategy and the effect of this strategy on security pricing; in consequence, the entrepreneur bears the agency costs of debt when he sells securities at t=0. Note here that we implicitly assume that outsiders have all the information about the firm's characteristics (in particular about the firm's investment opportunity set) and insiders' preferences (insiders' optimal private actions are common knowledge), which is necessary for them to rationally price their claims.

---

[6] Stapleton and Subrahmanyam (1984) show that in a discrete time framework, the combinations of negative exponential (CARA) utilities and normally distributed variables, or power (CRRA or HARA) utilities and lognormally distributed variables, yield a risk-neutral valuation relationship (RNVR) after a shift of the original parameters. In our model, the reader can consider that we look at the case where investors display constant absolute risk-aversion and returns are normal, but that all results are derived with parameters already shifted in such a way as to be able to use the tools of risk-neutral valuation. See also Ross (1978) for why we do not lose any generality assuming risk neutrality if markets are "reasonably complete".

[7] If one wants to relax this assumption, then all we need to assume is that if there were two different projects yielding θ and θ' respectively, all elements of the t=1 information set Φ_1 apart from X_1 are equally informative about θ and θ'.

A few comments on those assumptions are necessary. Assumption 2 may not be as innocuous as it seems. In our model, the outcome of the risky project may be arbitrarily negative. Since in the United States (as in most developed countries) equity has limited liability, we have to introduce a third claim. This claim is a purely negative one, borne by the government: if the firm faces hard times (i.e. a negative realization of θ), the government steps in to absorb the negative result (as, for instance, in the case of an environmental catastrophe where the state has to bear the cleanup costs)[8]. We will refer to the combined value of debt and equity as the "market value of the firm" since this is the amount for which the firm can be sold to the public at t=0. If we add to this the negative claim, we'll refer to it as the "social value of the firm".

Assumption 3 tells us that the firm chooses to issue outside claims (debt in our case) when it is common knowledge that there is no asymmetric information between the manager of the firm and outsiders. It is not clear whether it would be possible at all to issue further debt (no matter at what rate), when asymmetric information may lead to a "lemons" problem (see Akerlof (1970)). This is an issue we are considering in a companion paper.

Finally, Assumption 5 clarifies what is known by everybody and what is private information: apart from the realization x_1 that only the manager observes, everything else is publicly known; in particular: ρ_m, the speed of resolution of uncertainty for the particular firm/industry, $E(\theta)$, $Var(\theta)$, $E(X_1)$, $Var(X_1)$ and all other parameters, as well as the preferences of the manager, the firm's characteristics and the fact that the manager privately observes x_1, thus enabling bondholders to rationally price their claim. All other assumptions are standard and do not affect crucially our results.

The problem here is threefold. First of all, we would like to determine to what extent the informativeness of the signal observed by the manager before making the investment decision will affect her deviation from the optimal investment policy. Secondly, we set out to endogenously characterize the optimal contracts to limit such deviations. And finally, the resulting yield premium demanded on corporate debt has to be evaluated.

---

[8] Alternatively, this third claim could be an insurance policy for which the firm would have to pay a premium upfront. This complicates the model unduly, so we'll assume that the government steps in in case of a negative final cash flow but ensure in choosing our simulation parameters that the probability of this event happening is arbitrarily small.

3.   Solving the Problem

3.1. Some Preliminary Results

In the framework outlined above, it can be shown that the signal observed by outsiders at t=1 (defined as X_1 plus white noise) is also bivariate normally distributed with the outcome of the risky technology, although it is less informative about θ than X_1:

> ***Proposition 1***: *Y_1 and θ are bivariate normal random variables with correlation coefficient $\rho_b = \rho_m \rho_{XY}$ where ρ_{XY} is the correlation coefficient between X_1 and Y_1 (strictness of disclosure requirements) and ρ_m is the correlation coefficient between X_1 and θ (informativeness of manager's signal).*

Proof: $\rho_b \equiv \dfrac{\mathrm{cov}(\theta, Y_1)}{\sigma_\theta \sigma_{Y_1}} = \dfrac{\mathrm{cov}(\theta, X_1 + \varepsilon)}{\sigma_\theta \sigma_{Y_1}} = \rho_m \dfrac{\sigma_{X_1}}{\sigma_{Y_1}}$. Since $\rho_{XY} = \dfrac{\mathrm{cov}(X_1, X_1 + \varepsilon)}{\sigma_{X_1} \sigma_{Y_1}} = \dfrac{\sigma_{X_1}}{\sigma_{Y_1}}$, $\rho_b = \rho_m \rho_{XY}$.

Since $\mathrm{var}(Y_1 \mid X_1) = \sigma_\varepsilon^2 \equiv \sigma_{Y_1}^2 (1 - \rho_{XY}^2)$, Y_1 is a more noisy estimate of X_1 (larger σ_ε) if and only if ρ_{XY} is smaller. It becomes in turn a less reliable predictor of θ (smaller ρ_b) and leaves outsiders with a higher residual uncertainty $\sigma_{\theta|Y_1}^2 = \sigma_\theta^2 (1 - \rho_b^2)$. Hence, the larger $\rho_m^2 - \rho_b^2$ and σ_ε^2 or the smaller ρ_{XY}, the larger the extent of asymmetric information. In the sequel, we'll mainly do comparative statics with respect to ρ_{XY}, since we have a more intuitive feeling for it (see next paragraph), and since it does not need to be scaled (as opposed to σ_ε).

A useful way to think about it is as follows: ρ_m is the informativeness of the manager's signal and is the true speed of temporal resolution of uncertainty in a given firm/industry/market. It therefore gives us *inter-industry/market* comparative statics (once we have controlled for the intrinsic project risk σ_θ). However, *within* a given industry/market, disclosure requirements vary (due to size differences or the stock exchange on which a stock is traded) and therefore outsiders will be more or less able to monitor actions by insiders. Typically, large firms listed on the NYSE will have stricter disclosure requirements (larger ρ_{XY}) than smaller firms listed on the NASDAQ, leaving outsiders with a higher "signal precision" ρ_b[9]. This ρ_b *cannot*, however, be seen as a measure

---

[9] The term "disclosure requirements" encompasses only the *verifiable* information the manager releases. Big firms will probably enjoy fewer information asymmetries because more analysts follow them rather than due to significant differences in disclosure requirements between, say, the NYSE and the NASDAQ markets. Alternatively, our problem can be seen as looking at different firms within the same industry, but differing along two dimensions: the speed at which the uncertainty surrounding their operations is resolved (as measured by ρ_m) and the extent of information asymmetry between insiders and outsiders (measured by ρ_{XY}).

of the speed at which uncertainty is resolved, since it is the correlation coefficient between the return on the risky technology θ and a signal which can be more or less noisy *irrespective of the industry (or firm) considered*; ρ_b thus measures the speed of resolution of uncertainty "tainted" by asymmetric information. Outsiders, who know the value of ρ_m (but do not observe x_1), are aware that the ρ_b based on which they will make decisions is only the product of ρ_m and ρ_{XY} and compute true values for financial claims using ρ_m (which is common knowledge). In order to later understand the general solution to the problem, as well as notations, we will now present the case where $\rho_{XY}=0$. This corresponds to the case where outsiders are fed no information whatsoever (or an infinitely noisy, therefore irrelevant, information). We'll assume that outsiders do not write any contract *ex-ante* and will show in a later section, when we look at the general case, that they would not even if they could: contracting is suboptimal when the extent of asymmetric information is too large.

3.2. Example: the Opaque Firm

3.2.1. The Firm's Investment Policy

At t=1, the manager of a levered firm, subject to a managerial contract that perfectly aligns her incentives with shareholder interests, will seek to

$$\max_{Q \in [0,I]} \frac{1}{r_2} E\left[\left\{ Q\theta + (I-Q)r_2 - F \right\}^+ \mid X_1 = x_1 \right]$$

where $(w)^+$ stands for max(w,0). For an all-equity firm, the face value of the debt F has to be replaced by zero, in which case the objective function above represents the firm-value maximizing (first-best) investment policy.

More generally, the payoff to the shareholders is positive if and only if $Q\theta+(I-Q)r_2>F$ or $\theta \geq [F-(I-Q)r_2]/Q \equiv \theta^*$ (note here the circularity of the problem: θ^* will determine the shareholders' optimal Q, but Q enters the formula for θ^*). If the manager chooses to invest very little money in the risky technology (Q tends to 0), two possible cases arise: i) if $Ir_2 \geq F$, θ^* tends to $-\infty$: the firm will be solvent in all states of nature since its final wealth will be Ir_2, which is enough to cover its debt obligation of F; ii) if $Ir_2<F$, θ^* tends to $+\infty$: the firm will never be solvent since its final wealth, Ir_2, is insufficient to cover its debt obligation. Since no bank or individual would lend money with a face value greater than Ir_2 to our firm (which, in order to be solvent, would *have to* invest a positive Q in the risky technology, no matter how negative x_1 is, since investing in the risky technology is the only way to have a positive probability of avoiding bankruptcy), we shall assume that $F \leq Ir_2$. The fact that θ^* increases with F reflects the fact that as the amount of promised debt increases, the range of "states" where the firm defaults, $(-\infty,\theta^*]$, expands[10].

The problem then becomes, denoting $P(\theta|X_1)$ the probability distribution function of θ conditional on X_1,

---

[10] We call "state of the world" a particular x_1 if we are sitting at time 1, a particular θ or more generally a particular couple (x_1, θ) if we are sitting at t=2.

$$\max_{Q \in [0,I]} \int_{\theta^*}^{\infty} \left[Q\theta + (I-Q)r_2 - F \right] dP(\theta \mid X_1) \Leftrightarrow \max_{Q \in [0,I]} U(Q, \rho_m, x_1, F), \qquad (1)$$

where we have underscored the dependence of this indirect utility function on ρ_m and the particular realization x_1 observed at t=1. It is shown in Appendix 1 that this is equivalent to maximizing the following expression over $Q \in [0,I]$:

$$U(Q, \rho_m, x_1, F) = \left[QE(\theta \mid X_1 = x_1) + (I-Q)r_2 - F \right] \Phi(B_{x_1}) + Q\sqrt{\text{var}(\theta \mid X_1 = x_1)} \varphi(B_{x_1}) \qquad (2)$$

where $B_{x_1} \equiv \dfrac{E(\theta \mid X_1 = x_1) - \theta^*}{\sqrt{\text{var}(\theta \mid X_1 = x_1)}}$ and Φ and φ denote the standard normal cumulative probability function and

density function respectively.

The expression (2) is the expected value as of t=1 of the cash flows to be received by shareholders at t=2. It has to be discounted at the riskless rate r_2 to yield the share price at t=1, as a function of the quantity invested Q, the particular "state of the world" x_1 and the particular pattern of resolution of uncertainty (ρ_m shows up as well in $E(\theta|X_1=x_1)$ as in $\text{var}(\theta|X_1=x_1)$). Expression (2) can be interpreted as follows: the first term tells us that the shareholders get whatever is left of the cash flows to the firm after bondholders have been repaid if the firm is solvent (which happens with a probability of $\Phi(B_{x_1})$); the second term tells us that shareholders should be concerned not only with whether they finish "in-the-money", but by how much (the firm is now long an asset that bears a normal return and thus has to deal with this convexity feature).

It is shown in Appendix 1 that the function $U(Q,\rho_m,x_1,F)$ is convex, leading to a corner solution. Whether Q=0 or Q=I will depend on the particular x_1 observed by the manager: since $\lim_{x_1 \to -\infty, Q \to 0} \partial U(Q, \rho_m, x_1, F)/\partial Q = -\infty$, investing everything in riskless bills may yield a higher value for the objective function for low values of x_1. We now characterize the manager's investment policy.

3.2.2. Characterizing the Shareholder-Aligned Manager's Risk-Shifting Incentives

A risk-shifting (overinvestment) region is defined as a range of x_1 values for which a levered firm will invest in the risky venture whereas an all-equity firm will prefer to put the money into riskless bills, i.e. values of x_1 for which $U(I,\rho_m,x_1,F)>U(0,\rho_m,x_1,F)$ whereas $U(I,\rho_m,x_1,0) \leq U(0,\rho_m,x_1,0)$. The first task is to characterize the values of x_1 for which the firm decides to invest in the risky technology:

Lemma 1: *Given a certain pattern of temporal resolution of uncertainty ρ_m, there exists a unique cutoff value X_1^0 (resp. X_1^F) above which the manager of an all-equity (resp. levered) firm will invest in the risky technology.*

Proof: see Appendix 2.

For the sake of clarity, we will refer to a managerial policy investing in the risky technology for all $x_1 \geq X_1^F$ as investment policy $[X_1^F]$. Note that the investment policy $[X_1^0]$ is the one that could have been achieved if the realization x_1 had been perfectly observed by all parties and if a complete set of enforceable contracts specifying any investment policy could have been written.

Consistent with the extensive literature on agency games, it can be shown that the larger the amount of debt outstanding, the lower the investment cutoff X_1^F (i.e. the larger the x_1-region $[X_1^F, \infty]$ over which the firm decides to invest). In particular, a levered firm will never invest less than an all-equity one in the sense that the x_1-range that triggers investment by the levered firm contains the one that triggers investment by the all-equity firm. We therefore decide to call the quantity $X_1^0 - X_1^F$ the extent of risk-shifting, since it characterizes the region where signals are too low for the manager of an all-equity firm to invest, while the manager of a levered firm still chooses the risky technology. We summarize these different results in the following lemma:

Lemma 2:   i) *For a given ρ_m, the manager of a levered firm will overinvest compared to the manager of an all-equity firm in the sense that $X_1^F < X_1^0$; as a consequence, the investment policy $[X_1^F]$ gives rise to a t=2 cash flow distribution that is riskier, in the Rothschild and Stiglitz (1970, 1971) sense, than the t=2 cash flow distribution from the investment policy $[X_1^0]$.*
ii) The ex-ante probability of investing in the risky technology is increasing in the amount of outstanding debt F; in particular, it is greater for a levered firm than for an all-equity one.
iii) The extent of risk shifting is strictly increasing in the face value F of the firm's debt. Hence, the terminal cash flow distribution resulting from the investment policy $[X_1^F]$ is strictly riskier the larger F.

Proof: i) see Appendix 2 for a proof that the investment cutoff X_1^F is decreasing in F. In particular, $X_1^F < X_1^0$ for any positive amount of debt F. It then suffices to notice that over the range $[X_1^F, X_1^0]$, a strictly larger quantity will be invested in the risky technology by the levered firm, raising the variance of the final cash flows (while for any $x_1 \in [X_1^F, X_1^0]$, $E[\theta \mid X_1 = x_1] < E[\theta \mid X_1 = X_1^0]$).

ii) The ex-ante (t=0) probability that a firm with an investment policy $[X_1^F]$ will invest is $1 - \Phi((X_1^F - \overline{X}_1)/\sigma_{X_1})$, which is increasing in F since X_1^F is decreasing in F.

iii) Since X_1^0 doesn't depend on F, the quantity $X_1^0 - X_1^F$, previously shown to be positive for any positive F, is therefore increasing in F. The lower cutoff value X_1^F resulting from an increase in the amount of outstanding debt yields the added risk to the distribution of terminal cash flows. $\square$

Figure 2 gives a quick synopsis of the situation. As more debt is issued (i.e. as the probability of insolvency looms larger), the investment cutoff for the levered firm moves to the left, yielding a larger investment region $[X_1^F, \infty)$; in particular, the risk-shifting region $[X_1^F, X_1^0]$ widens: the incentive to increase the risk of the firm becomes more acute.

[Insert Figure 2 here]

3.2.3. Temporal Resolution of Uncertainty and Risk-Shifting Incentives

Our model therefore confirms that the distortion in the investment policy of a shareholder-aligned manager becomes more pronounced as the amount of debt outstanding increases. However, a far more interesting task is to find out how risk-shifting incentives vary with the pattern of temporal resolution of uncertainty ρ_m. Several results will be proven here.

Theorem 1: *i) The cutoffs X_1^0 and X_1^F are strictly increasing in ρ_m.*

ii) As a result, the risk of the firm's terminal cash flow distribution and the probability of investing in the risky technology are decreasing in ρ_m,

Proof: see Appendix 3 for a proof of i). As for ii), it follows directly from Lemma 2: if a larger ρ_m triggers a higher cutoff value X_1^0 (resp. X_1^F), then distribution of the terminal cash flows resulting from the investment policy $[X_1^0]$ (resp. $[X_1^F]$) implemented by the manager is strictly less risky and the probability of investing, $1 - \Phi((X_1^F - \bar{X}_1)/\sigma_{X_1})$, is lower.

This yields our first directly testable implication: firms operating in industries where uncertainty is resolved later will have more risky operations. Even when the intrinsic risk of a project (σ_0) is held constant (i.e. comparing two firms in the same industry), the firm for which uncertainty is resolved later will invest over a larger X_1-region than the firm for which uncertainty is resolved earlier. Hence, if a given firm decides to invest for a certain value of x_1, so will a firm facing more delayed uncertainty. This effect is independent of the fact that firms with later temporal resolution of uncertainty (lower ρ_m) may typically rely more on growth projects,

developing a new drug or software, with the risk that a competitor achieves it before them or that a regulatory agency prevents them from selling their product (shift in σ_θ).

Heuristically, this is due to the fact that when ρ_m is small, X_1 is less reliable in forecasting θ; hence the manager is ready to take the risk of investing in the risky project even for relatively low values of x_1, aware that the end result may still turn out to be very decent. The extent of the Bayesian update is small and the manager' beliefs are close to the prior ones, which would lead her to invest ($E(\theta)>r_2$). On the contrary, when ρ_m is larger, X_1 is more reliable in forecasting θ and a bad signal x_1 is a pretty good indication that the final outcome will be mediocre (the extent of the Bayesian update is larger).

It is interesting to notice that our result is *not* purely driven by the fact that the manager is faced with lower residual uncertainty $\sigma_\theta^2(1-\rho_m^2)$ when ρ_m gets larger. The investment cutoffs X_1^0 and X_1^F are not monotonic in σ_θ (see first simulation)! The reason is that σ_θ enters $E(\theta|X_1)$ as well. When x_1 is mediocre, i.e. in a region where it is not obvious whether the manager should invest or not, i.e. $x_1 < \overline{X}_1$ but x_1 is not too negative (we prove in Appendix 3 that both X_1^0 and X_1^F are negative), a larger σ_θ decreases $E(\theta|X_1=x_1)$, but at the same time "fattens" the distribution of θ conditional on x_1, therefore increasing the upside potential gains (upper tail of the distribution[11]). The investment cutoff is increasing in σ_θ for lower values of σ_θ (the larger ex-ante uncertainty mainly lowers $E(\theta|X_1=x_1)$ and results in less frequent investment). However over larger values of σ_θ, the potential gains over the (fatter) upper tail of the conditional density of θ outweigh the decrease in $E(\theta|X_1=x_1)$, yielding an investment cutoff that is decreasing in σ_θ: the likelihood that the manager will invest in the risky technology is increasing in the ex-ante uncertainty. This is especially true for larger amounts of debt since the cutoff X_1^F is implicitly defined as $E[\max(\theta, F/I)\,|\, X_1 = X_1^F] = r_2$. The larger F, the larger the proportion of the distribution of $E(\theta\,|\,X_1^F)$ that is discarded (lower tail) because of limited liability, and the larger the expectation over the remaining part of the truncated distribution.

The conclusion is different for a move in ρ_m. When ρ_m increases, it not only lowers the conditional expectation $E(\theta|X_1=x_1)$ for mediocre x_1's, it also makes the conditional distribution of θ tighter around this lower expectation: the manager is certainly *less* likely to invest when the signal is reliable and *bad*, with no upside potential. On the contrary, a lower value of ρ_m will both increase the conditional expectation $E(\theta|X_1=x_1)$ and the upside potential (fatter upper tail). Hence, given a certain ρ_m and σ_θ, a "borderline" x_1 will be unacceptable if we increase ρ_m but under some circumstances may be acceptable for a lower - or larger - σ_θ. We summarize the above discussion in Figure 3, but refer the reader to Epstein (1980) for more on why a larger ex-ante variance of

---

[11] The shareholder-aligned manager does not care about the lower tail because of limited liability.

the distribution of final cash flows and a more delayed resolution of uncertainty do not always yield the same *qualitative* results.

[Insert Figure 3 here]

The previous result holds for all-equity as well as levered firms. It turns out that it holds for the extent of debt-induced risk-shifting (the difference between X_1^0 and X_1^F, both of which increase with ρ_m) as well:

Theorem 2: *The extent of risk shifting* $X_1^0 - X_1^F$ *is strictly decreasing in the pattern of TRU* ρ_m.

Proof: see Appendix 4.

Note that our results are consistent with the existing research predicting that "companies whose value consists primarily of "growth options" as opposed to "assets in place" are likely to find debt financing very costly" (Myers (1977), p. 161). Intuitively "assets in place" are more numerous in industries where uncertainty is resolved earlier. This paper, however, explicitly uncovers the functional dependence between the pattern of temporal resolution of uncertainty and the investment policy distortions.

It is worth noting that the extent of risk-shifting is continuous as ρ_m tends to 1: observing x_1 fully reveals θ and whether the firm is levered or not, the risky project will be chosen if and only if $\theta(x_1) \equiv \overline{\theta} + \sigma_\theta (x_1 - \overline{X}_1)/\sigma_{X_1} \geq r_2$; any lower realization x_1, and the manager is better off investing in riskless bonds, securing a payoff of $Ir_2 - F$ for shareholders. However, this feature of continuity does not hold when ρ_m tends to 0. When $\rho_m = 0$, X_1 is irrelevant in forecasting θ and the managers of both the all-equity and the levered firm make a decision based on the unconditional mean, $E(\theta) > r_2$, and invest in the risky venture regardless of the particular realization x_1: there is no risk shifting. However, when ρ_m tends to 0, it is easily shown that the extent of risk shifting tends to infinity. Hence, there is a discontinuity in the extent of risk shifting in $\rho_m = 0$.

However, this does not have any impact on agency costs for two reasons: i) both X_1^0 and X_1^F tend to $-\infty$ as ρ_m approaches 0 and for all relevant purposes we can consider that both managers have the same investment policy (investing for all x_1's). It may be of more relevance to consider the *probability* of falling in this region, $\Phi((X_1^0 - \overline{X})/\sigma_{X_1}) - \Phi((X_1^F - \overline{X})/\sigma_{X_1})$, which indeed tends to 0; ii) when ρ_m is very low, X_1 is of very little relevance for predicting θ and even a significant distortion between the two investment policies would not yield significant *ex-ante* agency costs (formally, $E(\theta \mid X_1 = x_1) = \overline{\theta} + \rho_m \sigma_\theta (x_1 - \overline{X}_1)/\sigma_{X_1}$, and even if $x_1 - \overline{X}_1$ is very negative, $E(\theta|X_1 = x_1)$ will be close to $E(\theta)$ for very low ρ_m).

15

More generally, agency costs assume a bell-shaped curve as a function of ρ_m, with very low agency costs for ρ_m close to 0 (see explanation above) and for ρ_m close to 1 (because of the absence of investment distortion for large values of ρ_m). The largest agency costs are achieved for intermediate values of ρ_m, where the extent of risk-shifting is smaller than for lower values of ρ_m but causes much more harm since a mediocre x_1 is a more reliable signal that the final realization of θ will indeed be poor. For I=100, r_2=5%, E(θ)=15%, and σ_θ=0.6 (our standard scenario), agency costs represent up to 0.23% of firm value for F=50 (market value of debt/equity ratio: 0.675) and 1.16% for F=75 (market value of D/E ratio: 1.375). This is consistent with Leland (1998), who finds agency costs ranging from 0.32% to 1.22% of firm value depending on the risk management policy of the firm. The reason why the magnitude of agency costs is modest is that managers do not observe, as opposed to previous studies, any future parameter (e.g. future cash flows in case of success), but a signal that give them some *partial* information about future cash flows. Hence i) the magnitude of the asymmetric information is much smaller, and ii) if the policy conducted by the manager is much different from the one value-maximizers would choose, it does not ensure in any way that the final outcome of θ will make this deviation of consequence. Our agency costs are also rendered smaller by the fact that we do not consider the costs of financial distress (not only bankruptcy costs but also the impossibility to react freely to strategic moves by competitors if the firm is in the process of reorganization).

3.2.4. Temporal Resolution of Uncertainty and Corporate Debt Yields

Aware that the distribution of terminal cash flows will be less risky (Theorem 1) and that the extent of risk-shifting will be lower (Theorem 2), rational bondholders will demand a lower yield (or, equivalently, pay a larger price) for the bonds of a firm operating in a field where uncertainty is resolved earlier. We prove this in Appendix 5, after computing the t=0 value of bonds:

$$B_0(\rho_m, F, X_1^F) = \frac{1}{r_1 r_2}\left\{ F\Phi\left(\frac{X_1^F - \bar{X}}{\sigma_{X_1}}\right) + I\sigma_\theta (1-\rho_m^2)^{1/2} \int_{X_1^F}^{\infty} \left[A_{x_1}\Phi(A_{x_1}) + \varphi(A_{x_1}) - C_{x_1}\Phi(C_{x_1}) - \varphi(C_{x_1}) \right] dP(x_1) \right\}$$

where C_{x_1} stands for the function $\left[E(\theta | X_1 = x_1) - F/I\right]/\sigma_{\theta | X_1}$ (it is nothing else than the earlier function B_{x_1} when Q=I), $A_{x_1} = E(\theta | X_1 = x_1)/\sigma_{\theta | X_1} = C_{x_1} + F/(I\sigma_{\theta | X_1})$ and r_1 denotes the discount rate to be used from t=0 to t=1 (we would not lose any generality assuming that it is equal to 0).

> ***Theorem 3***: *The equilibrium prices of corporate bonds are increasing in ρ_m, the pattern of temporal resolution of uncertainty. Equivalently, the risk premium demanded on corporate bonds is decreasing in ρ_m.*

Note that this stems from two effects:

i) As ρ_m increases, so does the firm's market value because to the manager carrying out her investment policy based on a more reliable information[12]. Bondholders share this benefit with shareholders ("total firm value effect");

ii) as ρ_m increases, the extent of risk-shifting decreases (see Theorem 2), and bondholders benefit from a lower deviation from the socially optimal investment policy ("reduction in agency games effect").

Bondholders will demand lower bond yields from a firm with earlier resolution of uncertainty both because of a lower residual risk at the time of the investment decision ("total firm value effect") and because the manager will depart less from the value-maximizing investment policy. Both reasons follow from a later resolution of uncertainty. It however remains an open question whether TRU still has some power explaining the cross-sectional variation in yields demanded on corporate bonds, *once risk (default risk as well as overall business risk) has been controlled for*. A recent empirical investigation by Reisz and Perlich (1999) answers positively, with statistical as well as economic significance. This confirms the results of our simulations: keeping the business risk constant at $\sigma_\theta = 0.6$ and setting I=100, r_2=5% and E(θ)=15%, the yield premium is as large as 383 basis points for F=50 (market debt/equity ratio of 0.675) and 628 basis points for F=75 (market D/E ratio of 1.375). This is consistent with Leland (1998) and Duffie, Schroder and Skiadas (1996) for instance.

So far, outsiders were not allowed to contract. An unfortunate feature of our model is then that ρ_m measures as well the pattern of resolution of uncertainty as the extent of asymmetric information the manager enjoys at t=1 (how reliable the x_1 *observed only by her* is). Now, it is interesting to note that *the temporal resolution of*

---

[12] Following the same procedure as in the computation of bond prices, one can show that the levered firm's t=0 market value (value of debt + value of equity) is:

$$V_0(\rho_m, F, X_1^F) = \frac{I}{r_1 r_2}\left\{ r_2 \Phi\left(\frac{X_1^F - \overline{X}}{\sigma_{X_1}} \right) + \sigma_\theta (1-\rho_m^2)^{1/2} \int_{X_1^F}^{\infty} \left[A_{x_1}\Phi(A_{x_1}) + \varphi(A_{x_1}) \right] dP(x_1) \right\}$$

Totally differentiating this expression with respect to ρ_m (which appears directly as well as indirectly through X_1^F) yields a positive result (this differentiation is similar to the one in Appendix 5). One can see $\partial V_0(\rho_m, F, X_1^F)/\partial \rho_m > 0$ as the marginal increase in firm value that the choice by a marginally more informed the manager will bring (marginal value of information). If our model allowed the manager to buy more information for a certain cost, this increase in marginal value would have to be equated to the marginal cost of buying more information to find the optimal level of information.

Finally, since firm value is increasing in ρ_m, it is bounded below by $V_0(0, F, X_1^F) = V_0(0, 0, X_1^0) = IE[\max(\theta,0)]/r_1 r_2 =$

$I\left[\overline{\theta}\,\Phi(\overline{\theta}/\sigma_\theta) + \sigma_\theta \varphi(\overline{\theta}/\sigma_\theta) \right]/r_1 r_2 > I\overline{\theta}/r_1 r_2$, this last number being the social value of the firm for ρ_m=0. Since in our model E(θ)>r_2, this lower bound for the firm's value is strictly greater than I/r_1, the t=0 present value of the I dollars invested at t=1. The difference can be seen as the "growth options" of the firm (value of the opportunity to invest in a technology with E(θ)>r_2). Alternatively, it can be seen as the surplus of the entrepreneur when he sells the firm to outside shareholders and bondholders at t=0.

uncertainty effect overwhelms any opposite asymmetric information effect: if ρ_m represented only the extent of the information asymmetry between insiders and outsiders, rational bondholders would demand a larger yield for a *larger* ρ_m. Disentangling the two roles of ρ_m (TRU and information asymmetry) would hence not change the qualitative nature of our results. This is what we prove in the next section, where we allow outsiders to contract so as to mitigate the extent of risk-shifting.

3.3. The General Case: Outsiders Contract Optimally

3.3.1. Designing the Optimal Contractual Terms

In the more general case where outsiders observe as of t=1 a meaningful signal Y_1, it is in the entrepreneur's interest to write covenants as of t=0 when he sells claims to the firm's cash flows. Since he is the one who will bear the *ex-ante* agency costs of debt as he sells financial claims, his goal is to maximize the sale price of bonds plus the value of the equity he keeps. This is equivalent to saying that he will try to minimize the t=1 deviation from the firm value maximizing investment policy.

Since an investment policy can only be enforced based on verifiable information, outsiders will want to $\max_{Q \in [0,I]} \dfrac{1}{r_2} E\left[\left\{ Q\theta + (I-Q)r_2 \right\}^+ \mid Y_1 = y_1 \right]$ at t=1. This is achieved by investing in the risky technology if and only if $y_1 > Y_1^0$ defined, as in the previous section, by: $U(I, \rho_b, Y_1^0, 0) = U(0, \rho_b, Y_1^0, 0)$: at $y_1 = Y_1^0$, the firm as a whole is indifferent between investing in the risky technology or in riskless bonds. However, the decision is made here based on the observation of the signal Y_1, i.e. using the parameters $\rho_b = \rho_{XY}\rho_m < \rho_m$ and $\sigma_{Y_1}^2 = \sigma_{X_1}^2 + \sigma_\varepsilon^2$ instead of ρ_m and $\sigma_{X_1}^2$. Since the cutoff value is increasing in ρ (Theorem 1) and decreasing in the variance of the signal (i.e. σ_{X_1} for X_1^0, σ_{Y_1} for Y_1^0; see proof at the end of Appendix 3), $Y_1^0 < X_1^0$: the presence of the noise ε makes perfect contracting (i.e. enforcing the first-best investment policy) not possible.

The problem now is to figure out the terms of the bond covenants. Contracts based on *ex-post* realizations are assumed to be ruled out by legal structures. Since $\lim_{\rho_{XY} \to 0} Y_1^0 = \lim_{\rho_m \to 0} X_1^0 = -\infty$, forcing the manager to invest in the risky project if $y_1 > Y_1^0$ may in certain cases (i.e. for low values of ρ_{XY}) add very little value or even be counter-productive, creating new agency costs[13]. At first sight, the following contract could be written at t=0:

---

[13] This is similar to Myers' (1977) framework, where enforcing a minimum level of investment may create new agency costs due to forced investment in negative NPV states.

- if $Y_1^0 \geq X_1^F$, impose an investment cutoff of Y_1^0: if $y_1 \leq Y_1^0$, the manager will have to invest in riskless bonds and if $y_1 > Y_1^0$, the risky technology will be chosen.

- if $Y_1^0 < X_1^F$, bondholders will let the shareholder-aligned manager do as she pleases.

However, the same result can be achieved by merely stating that the manager is not allowed to invest below a certain cutoff Y_1^0, *regardless of how it compares to* X_1^{F} [14]. The reason is the following: insiders observe both realizations x_1 and ε_1 (the subscript on ε is added to denote a particular realization). Forcing the manager to invest above Y_1^0 (when $Y_1^0 \geq X_1^F$) leads her to invest in the risky technology if and only if $y_1 = x_1 + \varepsilon_1 > Y_1^0$, i.e. $x_1 > Y_1^0 - \varepsilon_1$. Now, in the case where $\varepsilon_1 > 0$ and Y_1^0 is only slightly larger than X_1^F, $Y_1^0 - \varepsilon_1 < X_1^F < Y_1^0$. For $x_1 \in [Y_1^0 - \varepsilon_1, X_1^F]$, neither shareholders nor bondholders would invest if complete contracting was possible (i.e. if outsiders could observe x_1) and *both are worse off if the covenant forces investment whenever* $y_1 > Y_1^0$, *even when* $Y_1^0 \geq X_1^F$. This problem is solved if the covenant only states that the manager cannot invest below a certain cutoff Y_1^0. If $Y_1^0 \geq X_1^F$, the interdiction of investing in the risky technology below Y_1^0 is binding (the shareholder-aligned manager would prefer to go down to X_1^F but is prevented from doing so by the bond covenants); if $Y_1^0 < X_1^F$, the manager would choose on her own not to invest below X_1^F. This leads us to:

> ***Proposition 2***: *The optimal covenant is one-sided, specifying that investment in the risky technology is forbidden for $y_1 \leq Y_1^0$.*

The extent of residual risk-shifting after this "partial" contracting is then $X_1^0 - \max(Y_1^0, X_1^F) = \min(X_1^0 - Y_1^0, X_1^0 - X_1^F)$, and in the sequel we will refer to optimal contracting as the bondholders imposing the interdiction to invest whenever their observation y_1 falls short of the cutoff Y_1^0, *no matter how the latter compares to* X_1^{F} [15]. Figure 4 shows the two possible cases, i.e. when contracting helps alleviating the manager's suboptimal investment policy and when bondholders cannot reduce the extent of risk-shifting.

[Insert Figure 4 here]

---

[14] It seems anyway that bond covenants usually do not specify what the manager has to do (e.g. above a certain cutoff) but merely state what she cannot do. We are grateful to Yaakov Amihud for pointing this out.

[15] It is still of interest to know under what parameters such a contracting is useful. Since X_1^F and Y_1^0 do not admit a closed-form representation, the particular ρ_m^* that equates X_1^F and Y_1^0 for a given ρ_{XY} will have to be solved for numerically.

New agency costs may be created by contracting under those terms. This is the case if Y_1^0 is very close to X_1^0 and $\varepsilon_1 < 0$. In that case, bondholders will rule out investing in the risky technology if $y_1 < Y_1^0$, i.e. if $x_1 < Y_1^0 - \varepsilon_1$. However, for $x_1 \in [X_1^0, Y_1^0 - \varepsilon_1]$, one should really invest, but the covenant will not allow it. This, however, is not a problem: Y_1^0 tends to X_1^0 only when ρ_{XY} tends to 1 (we'll shortly show that $\partial Y_1^0 / \partial \rho_{XY} > 0$; it should be obvious that $\lim_{\rho_{XY} \to 1} Y_1^0 = X_1^0$), in which case σ_ε tends to 0 ($\sigma_\varepsilon^2 = \sigma_{X_1}^2 (1 - \rho_{XY}^2) / \rho_{XY}^2$). The probability of a realization ε_1 such that $X_1^0 < Y_1^0 - \varepsilon_1$, $\Phi((Y_1^0 - X_1^0) / \sigma_\varepsilon)$, becomes in turn arbitrarily small.

The two following propositions will be useful shortly, when we analyze how the extent of risk-shifting and the yield demanded on corporate bonds can be reduced with optimal contracting

Proposition 3*: The benefit from contracting is non-decreasing in the face value of the debt outstanding.*

Proof: ◆ if $Y_1^0 \geq X_1^F$ for all $F \in [0, Ir_2]$, the improvement in terms of X_1-region due to contracting is $Y_1^0 - X_1^F$. Since $\partial Y_1^0 / \partial F = 0$ and $\partial X_1^F / \partial F < 0$, this improvement is strictly increasing in F.

◆ if $Y_1^0 < X_1^F$, the extent of residual risk-shifting is $X_1^0 - X_1^F$, the same as the original one (no benefit from contracting). However, as F increases, X_1^F decreases and at some point crosses to the left of Y_1^0 (since $\lim_{F \to Ir_2} X_1^F = -\infty$), yielding a positive improvement.

Since the precision of the signal is independent of how much debt the firm has issued[16] but risk-shifting becomes more severe the larger the amount of debt outstanding, there may be no improvement at all for relatively small amounts of debt (Y_1 is too noisy an estimate of X_1); however, as the proportion of debt in the capital structure increases, outsiders think twice before discarding the information y_1 (X_1^F crosses to the left of Y_1^0).

Proposition 4*: For a given firm/industry characterized by ρ_m, improvement in risk-shifting is non-decreasing in ρ_{XY}.*

---

[16] This might not be completely true: larger firms are likely to issue more debt, and at the same time, as seen previously, to have stricter disclosure requirements (or to be more closely followed by analysts). The precision of the signal is then increasing in the amount of debt issued. The same is true if more covenants are attached as the level of debt increases. We decide to ignore this for the time being.

Proof: $\partial(Y_1^0 - X_1^F)/\partial\rho_{XY} = \partial Y_1^0/\partial\rho_{XY} > 0$ (proof in Appendix 6). Therefore, if $Y_1^0 \geq X_1^F$ for a given

$F \in [0, Ir_2]$, the improvement increases with ρ_{XY}; if $Y_1^0 < X_1^F$, the improvement is independent

of ρ_{XY}, but as ρ_{XY} increases, so does Y_1^0 and at some point it crosses to the right of X_1^F,

yielding a positive improvement. The existence of a unique ρ_{XY}^* such that $Y_1^0 = X_1^F$ comes

from the facts that $\lim\limits_{\rho_{XY}\to 0} Y_1^0 = -\infty$ and $\lim\limits_{\rho_{XY}\to 1} Y_1^0 = X_1^0 > X_1^F$ (see Appendix 6), combined with

the Intermediate Value Theorem.

Ill-informed outsiders will be reluctant to contract, well aware that a mediocre y_1 may still be followed by a decent realization of θ. As Y_1 becomes more correlated with X_1, outsiders become more able (and therefore more willing) to effectively monitor the investment policy of the firm. Section 3.2. (where $\rho_{XY}=0$) can therefore be seen as dealing with the case where outsiders *choose* not to contract, given that it is impossible to improve on the manager's choice ($Y_1^0 < X_1^F$ for all ρ_m).

3.3.2. On the Extent of Residual Risk-Shifting After Optimal Contracting

We will here characterize the extent of the residual risk-shifting after optimal contracts have been written. The first step is to show that the benefit at t=1 from contracting, $\max(Y_1^0 - X_1^F, 0)$, is non-increasing in ρ_m:

> ***Theorem 4***: *at t=1, the benefit from contracting, i.e. the reduction in the extent of risk-shifting, is non-increasing in the pattern of resolution of uncertainty ρ_m. More precisely, for ρ_{XY} sufficiently large, i.e. for $\rho_{XY} \geq \underline{\rho_{XY}}$, there exists a unique ρ_m^* such that $Y_1^0 > X_1^F$ for $\rho_m < \rho_m^*$ and $Y_1^0 \leq X_1^F$ for $\rho_m \geq \rho_m^*$. The reduction in the extent of risk-shifting is then decreasing on $[0, \rho_m^*)$ and 0 on $[\rho_m^*, 1]$.*

> ***Proof***: see Appendix 6. If $\rho_{XY} < \underline{\rho_{XY}}$, we set $\rho_m^* = 0$: no industry benefits from contracting[17].

Contracting is more interesting for lower values of ρ_m since a levered firm operating in a field where uncertainty is resolved early does not have an investment policy significantly different from its all-equity equivalent. It therefore benefits less, for a fixed level of disclosure requirement ρ_{XY}, from contracting than a firm operating in a field where uncertainty is resolved later and which will have a tendency to risk-shift more significantly. Hence the following corollary:

---

[17] The firm becomes "opaque" and the reader is referred to Section 3.2.

Corollary 4: *Given a certain strictness of disclosure requirements $\rho_{XY} \geq \underline{\rho}_{XY}$, if a firm/industry benefits from contracting, so will a firm/industry for which uncertainty is resolved only later.*

Proof: this comes directly from Theorem 4 and the fact that for $\rho_{XY} \geq \underline{\rho}_{XY}$, it will be interesting for any firm below ρ_m^* (i.e. for which $Y_1^0 > X_1^F$) to contract. The decreasingness of $Y_1^0 - X_1^F$ in ρ_m over the region $[0, \rho_m^*]$ ensures that on this region, the lower ρ_m, the larger the reduction in the extent of risk-shifting.

Corollary 4 yields a ρ_{XY} / ρ_m^* "frontier" that will give us the ρ_m^* cutoff (below which firms benefit from contracting, above which contracting is useless) as a function of ρ_{XY}. Since ρ_m^* is shown in Appendix 6 to be increasing in ρ_{XY}, we get the following corollary as a bonus:

Corollary 4b: *The economy as a whole is better off, i.e. more industries benefit from contracting, the higher ρ_{XY}.*

It is pretty intuitive that the stricter the disclosure requirements, the more each firm/industry benefits from contracting. However, the preceding corollary also states that an increase in ρ_{XY} will bring *new* industries to contract; these are the industries that were just a little bit too mature (ρ_m too high) to benefit from a disclosure of quality ρ_{XY}, but who can gain from writing contracts based on information of quality $\rho_{XY}+\Delta\rho_{XY}$. In the limit, when ρ_{XY} tends to 1 (outsiders know as much as insiders), the optimal cutoff X_1^0 can be imposed by bondholders, leaving no residual agency costs: regardless of the amount of debt outstanding and the industry in which it operates, a firm will benefit from contracting. At the opposite end of the spectrum, when ρ_{XY} crosses $\underline{\rho}_{XY}$ to the left, no firm will benefit from contracting, no matter how large the amount of debt outstanding and regardless of the industry in which it operates.

However, for a given level of disclosure requirement ρ_{XY}, we are still left with the equivalent of Theorem 2, once optimal contracts have been written:

Theorem 2b: *The extent of residual risk-shifting $X_1^0 - \max(Y_1^0, X_1^F)$ is decreasing in the speed of resolution of uncertainty ρ_m.*

Proof: ♦ when $\rho_m \geq \rho_m^*$, $\partial(X_1^0 - X_1^F)/\partial\rho_m < 0$ as proved in Appendix 4;

22

◆ when $\rho_m < \rho_m^*$, we need to show that $\partial(X_1^0 - Y_1^0)/\partial\rho_m < 0$; since $X_1^0 = \lim_{\rho_{XY} \to 1} Y_1^0$, it suffices to show that $\partial^2 Y_1^0 /\partial\rho_{XY}\partial\rho_m < 0$, thereby ensuring that $\partial Y_1^0 /\partial\rho_m > \partial X_1^0 /\partial\rho_m$. This proof is very similar to the one in Appendix 4[18].

Partial contracting will thus mitigate risk-shifting, but will not destroy the monotonicity of the extent of risk-shifting in the pattern of temporal resolution of uncertainty. Thus all our discussion following Theorem 2 is still valid after bondholders contract as well as they can. We illustrate the case where $\rho_{XY} \geq \underline{\rho_{XY}}$ in Figure 5.

[Insert Figure 5 here]

3.3.3. On Corporate Debt Yields After Optimal Contracting

Our last task is to investigate how bond yields evolve with temporal resolution of uncertainty after bondholders impose the optimal investment cutoff $\max(X_1^F, Y_1^0)$. Bonds are priced as before (since, as noted earlier, ρ_m, which is public knowledge, will be used to price all assets), with the difference that the cutoff Y_1^0 is substituted for X_1^F whenever it is greater. It is shown in Appendix 7 that:

> ***Theorem 3b****: Once optimal contracts have been written, the equilibrium prices of corporate bonds are still increasing in ρ_m, the pattern of temporal resolution of uncertainty. Equivalently, the risk premium demanded on corporate bonds is still decreasing in ρ_m.*

As before, this stems from two effects: i) the "total firm effect" remains, i.e. the manager can carry out a more "educated" investment policy (alternatively, the residual variance of the risky project is lower), the larger ρ_m; bondholders share this benefit with shareholders[19] and ii) as ρ_m increases, the extent of risk-shifting $X_1^0 - \max(X_1^F, Y_1^0)$ decreases (see Theorem 2b), and bondholders benefit from a lower deviation from the socially optimal investment policy.

Finally, bond yields are non-increasing in the strictness of disclosure requirements (see Appendix 7):

---

[18] Since $\partial(Y_1^0 - X_1^F)/\partial\rho_m$ is negative for $\rho_m = \rho_m^*$ (i.e. when $Y_1^0 = X_1^F$), the residual risk-shifting, $\min(X_1^0 - Y_1^0, X_1^0 - X_1^F)$, is not differentiable in ρ_m at $\rho_m = \rho_m^*$; however, the reader should not be worried about non-differentiability on a set of measure zero.

[19] As noted in footnote 12, firm value is increasing in ρ_m as long as the cutoff the firm uses for investment is no larger than X_1^0. Since this is the case for both X_1^F and Y_1^0, it is also the case for $\max(X_1^F, Y_1^0)$ and firm value is still increasing in ρ_m.

23

Theorem 5*: Bond prices are non-decreasing in the strictness of disclosure requirements ρ_{XY} (non-increasing in the report noise σ_ε). Equivalently, the risk premium demanded on corporate bonds is non-increasing in ρ_{XY} (non-decreasing in the report noise[20]). More precisely, bond yields are decreasing in ρ_{XY} on $[0, \rho_m^*)$ and independent of ρ_{XY} on $[\rho_m^*, 1]$.*

Here, there is no "total firm value effect": the decrease in bond yields accompanying an increase in ρ_{XY} is solely due to the fact that the extent of residual risk shifting $X_1^0 - \max(X_1^F, Y_1^0)$ is decreasing in ρ_{XY} whenever $Y_1^0 > X_1^F$, i.e. when $\rho_m < \rho_m^*$ (following Proposition 4). This lower deviation from optimal investment policy enables bondholders to demand a lower yield premium.

This feature of our model is consistent with the existing literature on the effect of accounting reports on security prices. Most recently, Duffie and Lando (2001) refine the Leland (1994) model of default and allow for "imperfect information", i.e. outsiders observe only at discrete times a noisy accounting report on the value of assets, and survivorship. Their conclusion is that the zero-coupon credit spread is strictly increasing in the reporting noise level *a* (see their Figure 8). In our model, reports could be not only about accounting data, but also about the investment opportunity set facing the firm (how encouraging a t=1 marketing survey or a prototype test are). Since the quantity that characterizes the noisiness of this report in our model is σ_ε, our conclusion is the same as Duffie and Lando's one.

It is worth noting that share prices are decreasing in ρ_{XY} (see Appendix 7): although an increase in ρ_{XY} will increase firm value, this benefits bondholders and hurts shareholders, prevented from investing on the region $[Y_1^0, X_1^0]$. Equityholders rationally anticipate that their manager won't have total freedom to maximize the value of their claim and price it down accordingly. Note finally that we constrained ρ_{XY} to be the same across industries (hence our intuition that varying ρ_{XY} will yield *intra-industry* comparative statics). It would be of interest to let ρ_{XY} vary across industries. For instance, by positing that $\rho_{XY} = \rho_m^{n-1}$ or, equivalently, $\rho_b = \rho_m^n$ for n≥2, we are in effect imposing the restriction that there is more information asymmetry in industries where uncertainty is resolved later, which seems to be the intuitive case. We looked at the matter but did not report our results, which remain essentially unchanged.

---

[20] or, equivalently, non-decreasing in the amount of asymmetric information.

4.   Empirical Implications and Evidence

Our paper establishes that a shareholder-aligned manager will have more of an incentive to deviate from the value-maximizing investment policy the later the uncertainty surrounding the firm's operations is resolved. Rational bondholders will anticipate this behavior and demand larger yield premia on the debt of corporations with delayed resolution of uncertainty. Those managers will find it a simple matter to deviate significantly from the optimal investment policy; the results of such a deviation would take years to be noticed by creditors, by which time the value of the firm might be eroded beyond repair. Potential creditors will shy away from a situation they could neither monitor nor control and we would expect the level of debt in such an industry to be lower than in industries where uncertainty is resolved earlier (where the extent of risk-shifting is not only lower, as we have just proved, but also more effectively monitorable). This is reinforced by our Theorem 1, which states that the more delayed the uncertainty is, the larger the variance of distribution of terminal cash flows will be, which further decreases the debt capacity of the firm and increases risk premia.

The main problem will be of course to find a good proxy for the speed at which uncertainty is resolved. The difficulty here is to single out the effect of earlier or later temporal resolution of uncertainty from other effects: firms that have high R&D expenses or low quit ratios, i.e. firms we heuristically perceived as operating in a field where uncertainty is resolved later, are also firms

- that have more intangible assets and therefore larger expected bankruptcy costs (lower recovery rate on debt);
- with more risky operations and hence a larger probability of bankruptcy;
- the value of which is mainly accounted for by expected future earnings (their current earnings may even be negative), and therefore finding themselves in lower tax brackets (or even collecting tax loss carryforwards).

All these effects contribute in predicting lower debt ratios and higher bond yields for firms operating in fields where uncertainty is resolved later based only on existing theories of bankruptcy costs and tax considerations as determinants of capital structures. For instance, our conclusion that firms in a lower ρ_m field should rely less heavily on debt could be merely due to tax effects: progressivity in the tax structure implies that greater volatility in taxable income raises the firm's expected tax liabilities (see Smith and Stulz (1985)). Firms with later TRU (and hence riskier distributions of final cash flows) will have an incentive to reduce the amount of debt in their capital structure over the range of progressivity. Proxies related to the pattern of TRU that are used in the literature indeed control for other effects as well, such as the importance of growth options in the investment opportunity set of a particular firm; they are for instance the amount of R&D and advertising expenses (scaled by sales or market value of assets) as well as how job-specific the skills of the workers employed by the firm are. The empirical evidence listed along our paper's implications is to be taken with this in mind.

Bradley, Jarrell, and Kim (1984) regress the debt-to-value ratio on non-debt tax shields, R&D and advertising expenses *and firm volatility, thereby controlling for the overall volatility of assets as well as for the non debt tax shield due to R&D expenses*. The coefficient on R&D and advertising expenses is significantly negative (see their Table III), as our model predicted (firms that have high R&D expenses are usually firms in a late temporal resolution of uncertainty field and should therefore be more reluctant to issuing debt). They also report a "strong finding of intra-industry similarities in firm leverage ratios and of persistent inter-industry differences" in a way that supports our model: "54% of the cross-sectional variance in firm leverage ratios can be explained by industrial classification, with more variation in mean leverage ratios across industries than there is in firm leverage ratio within industries". Industries like "Drugs and Cosmetics", "Electronics" or "Petroleum Exploration", which we would heuristically characterize as having later temporal resolution of uncertainty, have significantly lower debt levels than industries like "Steel", "Telephone", "Electricity and Gas Utilities" or "Airlines", which seem to have earlier temporal resolution of uncertainty (see their Table I; their results are robust to the exclusion of regulated industries). A comparable test is carried out by Long and Malitz (1986) who look at 545 manufacturing firms, grouped into portfolios in order to hold the operating or business risk (as measured by unlevered betas) of all firms constant. The quartile with highest R&D expenses is the one with lowest leverage and the one with lowest R&D expenses is the one with highest leverage. The negative relationship between R&D (and advertising) expenses and leverage remains when the latter is regressed on the former, controlling for the firm systematic and residual risk and non-debt tax shields; these results carry over as well for portfolios as for individual firms.

Titman and Wessels (1988) use a factor-analytic technique, linear structural modeling, that mitigates the measurement problem encountered when working with proxy variables. They report that firms that are more "unique" tend to have lower debt/equity ratios. They define "uniqueness" as higher R&D and advertising expenses and lower quit ratios ("firms that sell products with close substitutes are likely to do less research and development since their innovations can be more easily duplicated" whereas "firms with relatively unique products are expected to advertise more and, in general, spend more in promoting and selling their products" and "employ workers with high levels of job-specific human capital who will thus find it costly to leave their jobs"[21]). Intuition dictates us that firms that operate in an industry where uncertainty is resolved late are, in a Titman and Wessels sense, more "unique"[22] and in that sense our model is supported by their empirical evidence (for evidence concerning the Japanese market, see Kester (1986) and Bronte (1982)).

---

[21] p. 5. They also control for the tax deductibility of R&D and some selling expenses. Titman (1984) also finds that "firms that can potentially impose high costs on their customers, workers and suppliers in the event of liquidation have lower debt ratios". This definition corresponds, once more heuristically, to our perception of a late temporal resolution of uncertainty pattern.

[22] As, for instance, computer or pharmaceutical firms, whereas firms that operate in a field where uncertainty is resolved earlier tend to offer products that are more substitutable and less characteristic of a given firm and to rely less heavily on human capital.

If firms operating in an environment where uncertainty is resolved late (low ρ_m) have already decided to take on debt (because insiders could not bring enough capital of their own or because the opportunity cost of inside funds or outside equity are even higher), we should see more representatives of creditors (bondholders or lending banks) sitting on the board of directors and more covenants in the bond contract-- so as to prevent shareholders from indulging too much in their risk shifting incentives -- than for firms operating in a higher ρ_m environment (see Theorem 4). Now, since the latter will typically have more risky debt outstanding, tests should control for it and investigate whether low ρ_m firms have more creditors on the board of directors and more debt covenants *per dollar lent*. Another solution to the risk-shifting problem may be to shorten debt maturity for those firms, rendering suboptimal investment by a shareholder-aligned manager more difficult. This is also in the best interest of the firms with most delayed resolution of uncertainty, since it may be able to refinance on better terms at the intermediate date. We therefore expect firms with more delayed resolution of uncertainty to issue shorter debt. Barclay and Smith (1995) finds supporting evidence in favor of this hypothesis (see also Stohs and Mauer (1996) and Guedes and Opler (1996)).

We finally address here the empirical implication about bond yields: bondholders, rationally anticipating risk-shifting incentives, will demand a higher yield from a corporation in a low ρ_m industry than from a corporation in a high ρ_m field (Theorems 3 and 3b). Here, some work has been done, investigating whether the pattern of temporal resolution of uncertainty has an effect on bond yields, *once the effect of classical determinants has been singled out*. Reisz and Perlich (1999) design three direct proxies for TRU, based on earnings forecasts and realizations. They are i) the time-series correlation between forecasted earnings and realized earnings, ii) the root mean square error of forecasts and ii) the dispersion in the earnings forecasts across analysts. These proxies try to estimate how much more difficult it is to forecast the long-term than it is to forecast the short term: a lumber company, say, will know almost as much about the longer-term as about the short-term, which is certainly not true for a company with delayed resolution of uncertainty. Reisz and Perlich (1999) find that indeed TRU still has some power in explaining cross-sectional variation in bonds yields once risk is controlled for. A firm in the last decile (late resolution) will have to offer up to 15.6 basis points more than a firm in the first decile (early resolution), *ceteris paribus*. Moreover, firms with more delayed resolution of uncertainty will have to offer *increasingly* larger yields the longer the bonds they issue: the maturity premium is monotonic in the pattern of TRU. They also find that although TRU is "priced" by the market, it is ignored by rating agencies in their grading process.

All of the above empirical implications should be mitigated when the existing regulation enforces verifiable disclosure requirements. In particular, this means that larger firms (or, more generally, firms listed on a stock

exchange enforcing stricter disclosure regulations) should display a lower extent of risk-shifting than smaller firms and should get away with lower yield premia[23].

As pointed out earlier, most of these empirical implications would have been achieved by "classical" corporate finance model based on bankruptcy and tax considerations. Moreover, most of the existing empirical evidence use proxies closely related to these factors and to the extent of growth opportunities in the firm's investment opportunity set. However, this does not undermine the validity of our theoretical model and its empirical implications: we come to the same conclusions as existing theories of bankruptcy costs and tax considerations, but *without even considering these effects, therefore in effect singling out the effect of temporal resolution of uncertainty on investment policies and bond yields*: we offer but one new element of explanation to the existing empirical evidence, to be added to the aforementioned consideration in further studies to acquire a fuller understanding of the available empirical evidence.

5.   Conclusion

In the tradition of papers dealing with the interactions between product markets and financial decisions, we argued that a firm's financial decisions *cannot be independent* of how quickly the uncertainty surrounding the firm's operations is resolved. Our paper's innovation is to show that firms for which uncertainty will be resolved later will suffer far more from the risk-shifting tendencies of a shareholder-aligned manager. As a result, the yield premium demanded on corporate bonds will be larger the later the uncertainty is resolved. This is mitigated when outsiders contract as well as they can based on the observation of a noisy signal, but the qualitative conclusions remain the same.

The fact that firm value and bond prices are larger the earlier the uncertainty is resolved does not, however, invalidate Ross' (1989) theory of "resolution irrelevancy". In our model, the difference in prices is due to the anticipation of different managerial behaviors that can lead to different cash flows. In Ross' words, changing the pattern of temporal resolution of uncertainty alters the no-arbitrage martingale pricing operator (which is equivalent to changing the state space spanning). As a result, risk-aversion is not even necessary for the speed of resolution of uncertainty to affect security prices.

We abstracted in our model from considering the different benefits of debt (tax deductibility of interest, disciplinary and signaling role of debt) as well as agency costs of equity. A more general model should include all of these in order to determine a firm's optimal capital structure (if it exists at all!). *However, these benefits and costs cannot be traded off against each other independently of the firm's pattern of temporal resolution of*

---

[23] With the problem that size may proxy for the liquidity of the firm's bonds as well.

uncertainty, upon which the optimal capital structure of the firm will depend. Further work should therefore investigate the optimal financing of different projects based on how quickly the uncertainty underlying their cash flows is resolved. This encompasses as well the nature of the claims to be used as their maturity: a richer menu of contractual forms for outside claims (for instance convertible debt) should be introduced. The model could be posited in a dynamic multiperiod framework and would investigate whether there is an optimal timing for raising money, in the sense of minimizing *ex-ante* agency costs as a function of the firm's pattern of TRU. Alternatively, it could try to explain the cross-sectional differences in dividend policies and add a "temporal" explanation to why mature firms differ markedly in this respect from growth firms.

Another avenue for future research is to investigate how the compensation package offered to the top officers of levered corporations should depend on the pattern of resolution of uncertainty surrounding the firm's operations[24]. Two contradicting intuitions are at play here: on the one hand, a manager should be compensated based on results. In a firm with more delayed resolution of uncertainty, this means delaying the manager's compensation, in the sense that more equity-like components (i.e. bonus, shares, and options) should be used. On the other hand, the compensation package should try to *disalign* the manager to mitigate the monotonic risk-shifting incentives we have uncovered in this paper (this will be easier in a market with substantial reputation costs for repeated suboptimal investment decisions). Hence, future research should not only add an element of explanation to the wide differences in managerial compensation across industries, but also suggest a way of driving managerial decisions closer to the optimal investment choices as a function of the industry in which the firm operates.

---

[24] Keeping in mind that the manager may not be shareholder-aligned in the first place. This may be due, for instance, to the fact that the executive's time horizon is relatively short while the value of the stock is the present value of dividends stretching to infinity.

Appendix

1. Computing the Objective Function and its First and Second Derivatives

We refer the reader to Epstein and Turnbull's (1980) Appendix 1, p.640, for a reminder on multivariate normal distributions. We'll just state the following result: if two scalar random variables X_1 and θ are jointly normally distributed, then the distribution of θ conditional on X_1 is also normal with the following parameters:

$$E(\theta \mid X_1 = x_1) = E(\theta) + \frac{\text{cov}(X_1, \theta)}{\text{var}(X_1)}(x_1 - E(X_1)) \equiv \bar{\theta} + \rho_m \frac{\sigma_\theta}{\sigma_{X_1}}(x_1 - \bar{X}_1)$$

$$\text{var}(\theta \mid X_1) = \sigma_\theta^2(1 - \rho_m^2) \equiv \sigma_{\theta \mid X_1}^2$$

where ρ_m stands for the correlation between X_1 and θ and σ_θ and σ_{X_1} are the respective standard deviations of θ and X_1. In the sequel we will write $E(\theta \mid X_1)$ for $E(\theta \mid X_1 = x_1)$, although it depends on a *particular realization x_1*. Equation (1) can be expressed as:

$$\int_{\theta^*}^{\infty} [Q\theta + (I - Q)r_2 - F] \frac{1}{\sqrt{2\pi\sigma_{\theta \mid X_1}^2}} \exp\left(-\frac{[\theta - E(\theta \mid X_1)]^2}{2\sigma_{\theta \mid X_1}^2}\right) d\theta$$

$$= [(I - Q)r_2 - F]\Phi\left(\frac{E(\theta \mid X_1) - \theta^*}{\sigma_{\theta \mid X_1}}\right) + Q\int_{\theta^*}^{\infty} \frac{\theta}{\sqrt{2\pi\sigma_{\theta \mid X_1}^2}} \exp\left(-\frac{[\theta - E(\theta \mid X_1)]^2}{2\sigma_{\theta \mid X_1}^2}\right) d\theta.$$

After a change of variable making $z \equiv [\theta - E(\theta \mid X_1)]/\sigma_{\theta \mid X_1}$ a standard normal variate, the second term is equal to

$$Q \int_{\frac{\theta^* - E(\theta \mid X_1)}{\sigma_{\theta \mid X_1}}}^{\infty} \left[(E(\theta \mid X_1) + z\sigma_{\theta \mid X_1}\right] \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = QE(\theta \mid X_1)\Phi\left(\frac{E(\theta \mid X_1) - \theta^*}{\sigma_{\theta \mid X_1}}\right) + Q\sigma_{\theta \mid X_1}\varphi\left(\frac{\theta^* - E(\theta \mid X_1)}{\sigma_{\theta \mid X_1}}\right).$$

Denoting $B_{x_1} \equiv \dfrac{E(\theta \mid X_1 = x_1) - \theta^*}{\sqrt{\text{var}(\theta \mid X_1 = x_1)}}$ (it is a function of a particular realization x_1, hence the superscript), a final

arrangement yields expression (2) in the text, from which

$$\frac{\partial U(Q, \rho_m, x_1, F)}{\partial Q} = [E(\theta \mid X_1) - r_2]\Phi(B_{x_1}) + [QE(\theta \mid X_1) + (I - Q)r_2 - F]\varphi(B_{x_1})\frac{\partial B_{x_1}}{\partial Q} + \sigma_{\theta \mid X_1}\varphi(B_{x_1}) + Q\sigma_{\theta \mid X_1}\frac{\partial \varphi(B_{x_1})}{\partial Q}$$

Now, $\dfrac{\partial \varphi(B_{x_1})}{\partial Q} = -B_{x_1}\varphi(B_{x_1})\dfrac{\partial B_{x_1}}{\partial Q}$ and $\dfrac{\partial B_{x_1}}{\partial Q} = \dfrac{F - Ir_2}{Q^2\sigma_{\theta \mid X_1}}$. Hence,

$$\frac{\partial U(Q, \rho_m, X_1, F)}{\partial Q} = [E(\theta \mid X_1) - r_2]\Phi(B_{x_1}) + \sigma_{\theta \mid X_1}\varphi(B_{x_1}) + \varphi(B_{x_1})\frac{F - Ir_2}{Q^2\sigma_{\theta \mid X_1}}[Ir_2 + Q[E(\theta X_1) - r_2] - F - Q\sigma_{\theta \mid X_1}B_{x_1}]$$

$$= [E(\theta \mid X_1) - r_2]\Phi(B_{x_1}) + \sigma_{\theta \mid X_1}\varphi(B_{x_1}) \tag{A1}$$

which can be negative for sufficiently low values of Q and x_1 ($\lim_{\substack{x_1 \to -\infty \\ Q \to 0}} B_{x_1} = +\infty$).

Differentiating this once more yields:

$$\frac{\partial^2 U(Q,\rho_m,x_1,F)}{\partial Q^2} = \varphi(B_{x_1})\frac{F-Ir_2}{Q^2\sigma_{\theta|X_1}}\left[E(\theta\,|\,X_1)-r_2-B_{x_1}\sigma_{\theta|X_1}\right] = \varphi(B_{x_1})\frac{(F-Ir_2)^2}{Q^3\sigma_{\theta|X_1}} > 0 \ \text{ for all } Q>0.$$

2. Proof of Lemmas 1 and 2

The manager of a levered firm compares $U(I,\rho_m,x_1,F)$ with $U(0,\rho_m,x_1,F)$ in deciding whether to invest or not. We therefore consider here the quantity $\Delta U(\rho_m,x_1,F)\equiv U(I,\rho_m,x_1,F)-U(0,\rho_m,x_1,F)$. Since $\lim_{Q\to 0} B_{x_1} = +\infty$ for all values of x_1, $\lim_{Q\to 0} U(Q,\rho_m,x_1,F) = Ir_2 - F$ (note the continuity of U when Q tends to 0: as we noted earlier, if Q=0, shareholders are left with Ir_2-F after investing everything in riskless Treasury bills and repaying bondholders) and

$$\Delta U(\rho_m,x_1,F) = \left[IE(\theta\,|\,X_1)-F\right]\Phi\left(\frac{IE(\theta\,|\,X_1)-F}{I\sigma_{\theta|X_1}}\right) + I\sigma_{\theta|X_1}\varphi\left(\frac{IE(\theta\,|\,X_1)-F}{I\sigma_{\theta|X_1}}\right) - Ir_2 + F$$

$$\equiv I\sigma_{\theta|X_1}\left[C_{x_1}\Phi(C_{x_1}) + \varphi(C_{x_1})\right] - Ir_2 + F \quad \text{where} \quad C_{x_1} \equiv \frac{IE(\theta\,|\,X_1=x_1)-F}{I\sigma_{\theta|X_1}} \qquad (A2)$$

<u>NB</u>: conditional on investing in the risky technology following a particular observation x_1, $\Phi(C_{x_1})$ is the t=1 probability of the final cash flow $I\theta$ exceeding the promised debt payment F, i.e. the probability of solvency

at t=2: $\int_{F/I}^{\infty} dP(\theta\,|\,X_1) = \Phi(C_{x_1})$.

Differentiating $\Delta U(\rho_m,x_1,F)$ with respect to x_1 yields:

$$\partial\Delta U(\rho_m,x_1,F)/\partial x_1 = I\sigma_{\theta|X_1}\Phi(C_{x_1})\partial C_{x_1}/\partial x_1 = I\rho_m\sigma_\theta\Phi(C_{x_1})/\sigma_{X_1} > 0 . \qquad (A3)$$

Therefore $\Delta U(\rho_m,x_1,F)$ is strictly increasing in x_1. Given that $\lim_{x_1\to -\infty}\Delta U(\rho_m,x_1,F) = -Ir_2 + F < 0$, $\lim_{x_1\to +\infty}\Delta U(\rho_m,x_1,F) = +\infty$ and the continuity of $\Delta U(\rho_m,x_1,F)$ in x_1, there is a unique x_1-cutoff X_1^F satisfying $\Delta U(\rho_m,X_1^F,F) = 0$ by the Intermediate Value Theorem. □

To prove that $X_1^0 > X_1^F$ for any positive amount of risky debt F, it suffices to show that X_1^F is decreasing in F. Hence any positive value of F will yield a cutoff X_1^F strictly smaller than X_1^0, and the larger the amount of debt outstanding, the lower the investment cutoff X_1^F will be. Now, X_1^F is defined as the value of x_1 for which $\Delta U(\rho_m,x_1,F)$=0. Keeping $\Delta U(\rho_m,X_1^F,F)$ equal to 0, implicit differentiation yields:

$$\frac{\partial X_1^F}{\partial F} = \left. \frac{-\dfrac{\partial \Delta U(\rho_m, x_1, F)}{\partial F}}{\dfrac{\partial \Delta U(\rho_m, x_1, F)}{\partial x_1}} \right|_{x_1 = X_1^F} = \left. -\frac{I\sigma_{\theta|X_1} \dfrac{\partial C^{x_1}}{\partial F} \Phi(C_{x_1}) + 1}{I\rho_m \sigma_\theta \Phi(C_{x_1}) / \sigma_{X_1}} \right|_{x_1 = X_1^F} = \frac{\Phi(C_{X_1^F}) - 1}{I\rho_m \sigma_\theta \Phi(C_{X_1^F}) / \sigma_{X_1}} < 0 \qquad \text{(A4)}$$

where $C_{X_1^F}$ denotes the expression C_{x_1} taken in $x_1 = X_1^F$. $\square$

3. Proof of Theorem 1

We will prove here that the earlier the temporal resolution of uncertainty (i.e. the larger ρ_m), the larger the cutoff value X_1^0 (resp. X_1^F) used by the manager of an all-equity (resp. levered) firm. Following the same argument as before, we keep $\Delta U(\rho_m, X_1^F, F)$ equal to 0 and use implicit differentiation. Now, since

$$\frac{\partial C_{x_1}}{\partial \rho_m} = \frac{x_1 - \bar{X}_1}{\sigma_{X_1} \left(1 - \rho_m^2\right)^{3/2}} + \frac{\rho_m \left[I\bar{\theta} - F\right]}{I\sigma_\theta \left(1 - \rho_m^2\right)^{3/2}}, \qquad \text{(A5)}$$

$$\frac{\partial \Delta U(\rho_m, x_1, F)}{\partial \rho_m} = -I\sigma_\theta \rho_m \left(1 - \rho_m^2\right)^{-1/2} \left[C_{x_1} \Phi(C_{x_1}) + \varphi(C_{x_1})\right] + I\sigma_\theta \left(1 - \rho_m^2\right)^{1/2} \left[\frac{\partial C_{x_1}}{\partial \rho_m} \Phi(C_{x_1})\right]$$

$$= \frac{I\sigma_\theta \left(x_1 - \bar{X}_1\right)}{\sigma_{X_1}} \Phi(C_{x_1}) - \frac{I\sigma_\theta \rho_m}{\left(1 - \rho_m^2\right)^{1/2}} \varphi(C_{x_1}) \qquad \text{(A6)}$$

and combining equations (A3) and (A6) we get:

$$\frac{\partial X_1^F}{\partial \rho_m} = -\frac{\dfrac{I\sigma_\theta \left(X_1^F - \bar{X}_1\right)}{\sigma_{X_1}} \Phi(C_{X_1^F}) - \dfrac{I\sigma_\theta \rho_m}{\left(1 - \rho_m^2\right)^{1/2}} \varphi(C_{X_1^F})}{I\rho_m \dfrac{\sigma_\theta}{\sigma_{X_1}} \Phi(C_{X_1^F})} = \frac{\sigma_{X_1}}{\left(1 - \rho_m^2\right)^{1/2}} \cdot \frac{\varphi(C_{X_1^F})}{\Phi(C_{X_1^F})} - \frac{(X_1^F - \bar{X}_1)}{\rho_m} \qquad \text{(A7)}.$$

To prove that this expression is positive, it suffices to show that $X_1^F < \bar{X}_1$ for all values of F. For that purpose, we introduce the quantity X_1^* such that $E(\theta \mid X_1 = X_1^*) = r_2$: $X_1^* \equiv \bar{X}_1 + (r_2 - \bar{\theta})\sigma_{X_1} / (\rho_m \sigma_\theta) < \bar{X}_1$ since E(θ)>r_2.

Now, $\partial U(Q, \rho_m, X_1^*, F) / \partial Q = \sigma_{\theta|X_1} \; \varphi((Ir_2 - F)/Q\sigma_{\theta|X_1})$ (see equation (A1)), which is positive for all Q$\in$[0,I].

Therefore $\Delta U(\rho_m, X_1^*, F) \equiv U(I, \rho_m, X_1^*, F) - U(0, \rho_m, X_1^*, F) > 0 \equiv \Delta U(\rho_m, X_1^F, F)$. The strict increasingness of ΔU in X_1 ensures that $X_1^* > X_1^F$ for all F$\in$[0,Ir$_2$). Hence $X_1^F < X_1^* < \bar{X}_1$ and the expression (A7) becomes positive for all F$\in$[0,Ir$_2$). $\square$

It is of interest to say one more word on X_1^*. It is the cutoff used by a manager who tries to

$$\max_{Q \in [0,I]} \frac{1}{r_2} E\left[\left\{Q\theta + (I - Q)r_2\right\} \mid X_1 = x_1\right]$$ (social value of the firm, including the responsibility to face a θ<0) and

would therefore invest whenever $E(\theta|X_1) \geq r_2$. Since $X_1^* > X_1^F$ for all $F \in [0, Ir_2)$, in particular for $F=0$, the manager of our all-equity firm still overinvests compared to the social optimum due to limited liability.

Finally, we can prove, using the same steps as above, that $\partial X_1^F / \partial \sigma_{X_1} = (X_1^F - \bar{X}_1) / [I(1 - \rho_m^2)^{1/2} \sigma_{X_1} \sigma_\theta \Phi(C_{X_1^F})]$, which is negative since $X_1^F < \bar{X}_1$. This will be useful when we consider outside contracts based on the signal Y_1.

4. Proof of Theorem 2

We will prove here that the earlier uncertainty is resolved (i.e. the larger ρ_m), the lower the extent of risk shifting $X_1^0 - X_1^F$. To show that $\partial(X_1^0 - X_1^F) / \partial \rho_m < 0$, i.e. $\partial X_1^0 / \partial \rho_m < \partial X_1^F / \partial \rho_m$ for all $F \in [0, Ir_2)$, it suffices to prove that $\partial^2 X_1^F / \partial \rho_m \partial F > 0$.

Now, drawing from the results in the previous appendix and differentiating (A7) with respect to F (remembering that $C_{X_1^F}$ depends on F both directly and through X_1^F),

$$\frac{\partial^2 X_1^F}{\partial \rho_m \partial F} = \frac{-\sigma_{X_1} \varphi(C_{X_1^F}) D \left[C_{X_1^F} \Phi(C_{X_1^F}) + \varphi(C_{X_1^F}) \right]}{(1 - \rho_m^2)^{1/2} \left[\Phi(C_{X_1^F}) \right]^2} - \frac{\Phi(C_{X_1^F}) - 1}{I \rho_m^2 \frac{\sigma_\theta}{\sigma_{X_1}} \Phi(C_{X_1^F})}$$

where we took the expression for $\partial X_1^F / \partial F$ from (A4) and

$$D \equiv \left[\frac{\partial C_{X_1^F}}{\partial F} + \frac{\partial C_{X_1^F}}{\partial X_1^F} \cdot \frac{\partial X_1^F}{\partial F} \right] = \frac{-1}{I \sigma_\theta (1 - \rho_m^2)^{1/2}} + \frac{\rho_m}{\sigma_{X_1} (1 - \rho_m^2)^{1/2}} \frac{\Phi(C_{X_1^F}) - 1}{I \rho_m \frac{\sigma_\theta}{\sigma_{X_1}} \Phi(C_{X_1^F})} = \frac{-1}{I \sigma_\theta (1 - \rho_m^2)^{1/2} \Phi(C_{X_1^F})}$$

Therefore

$$\frac{\partial^2 X_1^F}{\partial \rho_m \partial F} = \frac{\rho_m^2 \varphi(C_{X_1^F}) \left[C_{X_1^F} \Phi(C_{X_1^F}) + \varphi(C_{X_1^F}) \right] + (1 - \rho_m^2) \left[1 - \Phi(C_{X_1^F}) \right] \left[\Phi(C_{X_1^F}) \right]^2}{\rho_m^2 \frac{I \sigma_\theta}{\sigma_{X_1}} (1 - \rho_m^2) \left[\Phi(C_{X_1^F}) \right]^3}$$

It now suffices to remember that, by definition of X_1^F, $C_{X_1^F} \Phi(C_{X_1^F}) + \varphi(C_{X_1^F}) = \frac{Ir_2 - F_2}{I \sigma_\theta (1 - \rho^2)^{1/2}} > 0$ (see (A2)) to see that the expression above is positive. □

5. Temporal Resolution of Uncertainty and Corporate Debt Yields

We will here price the bonds of the firm in terms of an arbitrary X_1-cutoff ξ for the sake of generality (when considering partial contracting, we'll use an X_1-cutoff different from X_1^F).

◆ As of t=1, if $x_1<\xi$, the manager buys riskless bills and the bond will be worth F at t=2 (since $Ir_2>F$); if $x_1\geq\xi$, the manager invests in the risky technology and the bond will be worth $\max[0,\min(I\theta,F)]$ [25] at t=2. So conditional on investing in riskless bonds, the t=1 value of the bond is F/r_2. Conditional of investing in the risky technology, the value of the bond as of t=1, denoted $B_1(\rho_m,F,x_1)$, is:

$$B_1(\rho_m,F,x_1 \mid x_1>\xi) = \frac{1}{r_2}\left\{ F\int_{F/I}^{\infty} dP(\theta\mid X_1)+I\int_0^{F/I}\theta\, dP(\theta\mid X_1)\right\}$$

$$= \frac{1}{r_2}\left\{ F\Phi(C_{x_1})+IE(\theta\mid X_1)\left[\Phi(A_{x_1})-\Phi(C_{x_1})\right]+I\sigma_{\theta\mid X_1}\left[\varphi(A_{x_1})-\varphi(C_{x_1})\right]\right\}$$

$$= \frac{I\sigma_{\theta\mid X_1}}{r_2}\left[A_{x_1}\Phi(A_{x_1})+\varphi(A_{x_1})-C_{x_1}\Phi(C_{x_1})-\varphi(C_{x_1})\right]$$

where $A_{x_1} = \dfrac{E(\theta\mid X_1=x_1)}{\sigma_{\theta\mid X_1}} = C_{x_1}+F/(I\sigma_{\theta\mid X_1})>C_{x_1}$ for all x_1; once more, it is obvious from the derivation that $\Phi(C_{x_1})$ is nothing else than the t=1 probability of being solvent and that $\Phi(A_{x_1})$ is the probability of θ being positive, both conditional on investing in the risky technology.

NB: note that the next-to-last line in the expression for $B_1(\rho_m,F,x_1)$ tells us that at t=1 bondholders expect the firm to be solvent (and to receive F) with probability $\Phi(C_{x_1})$ and believe that with a probability of $\Phi(A_{x_1})-\Phi(C_{x_1})$ the firm will not be solvent but still have positive value, in which case they will take over, receiving on average $IE(\theta\mid X_1=x_1)$; the last term is a convexity adjustment.

◆ So as of t=0, the bond price is equal to:

$$B_0(\rho_m,F,\xi) = \frac{1}{r_1 r_2}\left\{ F\Phi\left(\frac{\xi-\bar{X}}{\sigma_{X_1}}\right)+I\sigma_\theta(1-\rho_m^2)^{1/2}\int_\xi^{\infty}\left[A_{x_1}\Phi(A_{x_1})+\varphi(A_{x_1})-C_{x_1}\Phi(C_{x_1})-\varphi(C_{x_1})\right]dP(x_1)\right\} \quad \text{(A8)}$$

as stated in the text.

Since the debt has the form of a zero-coupon bond, we define the per period yield premium demanded on the debt as $\sqrt{\dfrac{F}{r_1 r_2 B_0(\rho_m,F,\xi)}}$ -1 (bonds in our model have a maturity of two periods). Proving that this yield

---

[25] Since the return on the risky technology is normally distributed, it can assume an arbitrarily negative value. As stated in the text, we consider in that case that the society at large has to bear the consequences of a negative realization of θ (product liability suit, environmental catastrophe etc.). However, we can choose the parameters so as to make the probability of a negative cash flow, conditional on having invested in the risky technology, $\left(\int_{X_1^F}^{\infty}\left(\int_{-\infty}^{0} dP(\theta\mid X_1=x_1)\right)dP(x_1)\right)$ arbitrarily small.

With our parameters ($E(\theta)=1.15$, $\sigma_\theta=0.6$), it is never larger than 2.76% for any value of ρ_m.

premium is decreasing in ρ_m is therefore equivalent to proving that bond prices are increasing in ρ_m. Now, from (A8),

$$\frac{\partial r_1 r_2 B_0(\rho_m, F, \xi)}{\partial \rho_m} = \frac{F}{\sigma_{X_1}} \varphi\left(\frac{\xi - \bar{X}_1}{\sigma_{X_1}}\right) \frac{\partial \xi}{\partial \rho_m} - I\sigma_\theta \rho_m \left(1 - \rho_m^2\right)^{-1/2} \int_\xi^\infty \left[A_{x_1}\Phi(A_{x_1}) + \varphi(A_{x_1}) - C_{x_1}\Phi(C_{x_1}) - \varphi(C_{x_1})\right] dP(x_1)$$

$$+ I\sigma_\theta \left(1 - \rho_m^2\right)^{1/2} \left[\int_\xi^\infty \frac{\partial\left[A_{x_1}\Phi(A_{x_1}) + \varphi(A_{x_1})\right]}{\partial \rho_m} dP(x_1) - \left[A_\xi\Phi(A_\xi) + \varphi(A_\xi)\right]\varphi_{X_1}(\xi)\frac{\partial \xi}{\partial \rho_m}\right]$$

$$- I\sigma_\theta \left(1 - \rho_m^2\right)^{1/2} \left[\int_\xi^\infty \frac{\partial\left[C_{x_1}\Phi(C_{x_1}) + \varphi(C_{x_1})\right]}{\partial \rho_m} dP(x_1) - \left[C_\xi\Phi(C_\xi) + \varphi(C_\xi)\right]\varphi_{X_1}(\xi)\frac{\partial \xi}{\partial \rho_m}\right] \qquad \text{(A9)}$$

where $\varphi_{X_1}(\xi) = \dfrac{1}{\sqrt{2\pi}\sigma_{X_1}} e^{-\frac{1}{2}\left(\frac{\xi - \bar{X}}{\sigma_{X_1}}\right)^2}$ denotes the probability density function of X_1 taken in $x_1 = \xi$ and, as before,

the subscript ξ on A and C means that these expressions are also taken in $x_1 = \xi$. Here, we need to come back to our particular case where the cutoff is X_1^F and remember that

- $I\sigma_\theta\left(1 - \rho_m^2\right)^{1/2}\left[C_{X_1^F}\Phi(C_{X_1^F}) + \varphi(C_{X_1^F})\right] \equiv Ir_2 - F$;

- the functions $f: x \mapsto x\Phi(x) + \varphi(x)$ and $g: x \mapsto E(\theta \mid X = x)$ are increasing in x; hence

  $A_{X_1^F} \equiv \dfrac{E(\theta \mid X_1 = X_1^F)}{\sigma_{\theta|X_1}} < \dfrac{E(\theta \mid X_1 = X_1^0)}{\sigma_{\theta|X_1}} \equiv A_{X_1^0}$ since X_1^F is no larger than X_1^0 by Lemma 2 and

  $A_{X_1^F}\Phi(A_{X_1^F}) + \varphi(A_{X_1^F}) < A_{X_1^0}\Phi(A_{X_1^0}) + \varphi(A_{X_1^0}) \equiv \dfrac{r_2}{\sigma_\theta(1 - \rho_m^2)^{1/2}}$;

- $\varphi_{X_1}(X_1^F) = \varphi\left((X_1^F - \bar{X}_1)/\sigma_{X_1}\right)/\sigma_{X_1}$; and $\partial X_1^F/\partial \rho_m$;

and the derivative (A9) becomes strictly greater than:

$$-I\sigma_\theta\left(1 - \rho_m^2\right)^{-1/2} \int_{X_1^F}^\infty \left[\rho_m A_{x_1}\Phi(A_{x_1}) + \rho_m\varphi(A_{x_1}) - \left(1 - \rho_m^2\right)\Phi(A_{x_1})\frac{\partial A_{x_1}}{\partial \rho_m}\right] dP(x_1)$$

$$+I\sigma_\theta\left(1 - \rho_m^2\right)^{-1/2} \int_{X_1^F}^\infty \left[\rho_m C_{x_1}\Phi(C_{x_1}) + \rho_m\varphi(C_{x_1}) - \left(1 - \rho_m^2\right)\Phi(C_{x_1})\frac{\partial C_{x_1}}{\partial \rho_m}\right] dP(x_1)$$

$$= I\rho_m\sigma_\theta\left(1 - \rho_m^2\right)^{-1/2} \int_{X_1^F}^\infty \left[\varphi(C_{x_1}) - \varphi(A_{x_1})\right] dP(X_1) - I\sigma_\theta \int_{X_1^F}^\infty \frac{(X_1 - \bar{X}_1)}{\sigma_{X_1}}\left[\Phi(C_{x_1}) - \Phi(A_{x_1})\right] dP(x_1)$$

$$= I\rho_m\sigma_\theta\left(1 - \rho_m^2\right)^{-1/2} \int_{\frac{X_1^F - \bar{X}_1}{\sigma_{X_1}}}^\infty \left[\varphi(a_C + b_C z) - \varphi(a_A + b_A z)\right]\varphi(z)dz - I\sigma_\theta \int_{\frac{X_1^F - \bar{X}_1}{\sigma_{X_1}}}^\infty \left[\Phi(a_C + b_C z) - \Phi(a_A + b_A z)\right] z\varphi(z)dz$$

where we used the expression (A5) for $\partial C_{x_1}/\partial \rho_m$ and $\partial A_{x_1}/\partial \rho_m$ and $a_A = \bar{\theta}/\left[\sigma_\theta(1 - \rho_m^2)^{1/2}\right]$,

$a_C = a_A - F/\left[I\sigma_\theta(1 - \rho_m^2)^{1/2}\right]$, $b_A = b_C = \rho_m/(1 - \rho_m^2)^{1/2}$ are all constants and z is a standard normal variate.

Using equations (15) and (26) from Carr and Rubinstein (1995) to integrate the above normals and simplifying, the above expression is equal to $I\sigma_\theta \left[\Phi(A_{X_1^F}) - \Phi(C_{X_1^F}) \right] \varphi\left((X_1^F - \bar{X}_1)/\sigma_{X_1} \right) > 0$. $\square$

Note that when ρ_m tends to 1, the manager of a levered firm adopts the same cutoff policy as the manager of an all-equity firm. It can be shown that $\lim\limits_{\rho_m \to 1} X_1^F = \lim\limits_{\rho_m \to 1} X_1^0 = \lim\limits_{\rho_m \to 1} X_1^* = \bar{X}_1 + (r_2 - \bar{\theta})\sigma_{X_1}/\sigma_\theta$ and since $Ir_2 > F$, C_{x_1} tends to infinity for all x_1 greater than X_1^F. Hence $C_{x_1}\Phi(C_{x_1}) + \varphi(C_{x_1}) \sim C_{x_1}$, $A_{x_1}\Phi(A_{x_1}) + \varphi(A_{x_1}) \sim A_{x_1}$ and

$$\lim_{\rho_m \to 1} \left\{ I\sigma_{\theta|X_1} \int_{X_1^F}^{\infty} \left[A_{x_1}\Phi(A_{x_1}) + \varphi(A_{x_1}) - C_{x_1}\Phi(C_{x_1}) + \varphi(C_{x_1}) \right] dP(x_1) \right\}$$

$$= \lim_{\rho_m \to 1} \left\{ I\sigma_\theta \left(1 - \rho_m^2\right)^{1/2} \int_{X_1^F}^{\infty} \left[\frac{E(\theta|X_1)}{\sigma_\theta\left(1-\rho_m^2\right)^{1/2}} - \frac{E(\theta|X_1) - F/I}{\sigma_\theta\left(1-\rho_m^2\right)^{1/2}} \right] dP(x_1) \right\} = F\left[1 - \Phi\left(\frac{X_1^F - \bar{X}}{\sigma_{X_1}} \right) \right]$$

Inserting the result above in equation (A8) yields the result that $\lim\limits_{\rho_m \to 1} B_0(\rho_m, F, X_1^F) = F/(r_1 r_2)$: bond prices are continuous in $\rho_m = 1$. To see this, we just have to remember that when $\rho_m = 1$, θ is revealed as of t=1 and the manager invests in the risky project if and only if this already revealed θ is lager than r_2. Since $Ir_2 > F$, the firm will always be solvent; the promised payment can be discounted at the riskless rate to determine the bond price.

Finally, in much the same way as we did for bond prices, it can be shown that share price as well as the value of the claim held (unwillingly) by taxpayers are increasing in ρ_m: an earlier resolution of uncertainty is Pareto-improving.

6. Mitigating Risk-Shifting through Partial Contracting

Our first task is to prove Proposition 4, i.e. $\partial Y_1^0 / \partial \rho_{XY} > 0$. This is done in much the same way as in Appendix 3 and we'll go quickly through the proof. We keep $\Delta U(\rho_b, Y_1^0, 0) \equiv U(I, \rho_m, Y_1^0, 0) - U(0, \rho_m, Y_1^0, 0)$ constant at 0 (this is the implicit definition of Y_1^0, using the same notation as earlier), and use total differentiation. Now, $\partial \Delta U(\rho_b, y_1, 0)/\partial y_1 = I\rho_m \rho_{XY}^2 \sigma_\theta \Phi(A_{y_1})/\sigma_{X_1}$, where A_{y_1} denotes the expression $E(\theta|Y_1 = y_1)/\sigma_{\theta|Y_1}$ and we used the fact that $E(\theta|Y_1=y_1)$ depends on ρ_{XY} through both $\rho_b = \rho_{XY}\rho_m$ and $\sigma_{Y_1} = \sigma_{X_1}/\rho_{XY}$. Finally, since

$$\frac{\partial A_{y_1}}{\partial \rho_{XY}} = \frac{\rho_m \rho_{XY}\sigma_\theta(y_1 - \bar{X}_1)(2 - \rho_m^2\rho_{XY}^2)/\sigma_{X_1} + \rho_m^2\rho_{XY}\bar{\theta}}{\sigma_\theta\left(1 - \rho_m^2\rho_{XY}^2\right)^{3/2}},$$

$$\frac{\partial \Delta U(\rho_b, y_1, 0)}{\partial \rho_{XY}} = -I\sigma_\theta \rho_m^2 \rho_{XY}(1-\rho_m^2 \rho_{XY}^2)^{-1/2}\left[A_{y_1}\Phi(A_{y_1})+\varphi(A_{y_1})\right]+I\sigma_\theta(1-\rho_m^2\rho_{XY}^2)^{1/2}\Phi(A_{y_1})\frac{\partial A_{y_1}}{\partial \rho_{XY}}$$

$$= 2I\rho_m\rho_{XY}\frac{\sigma_\theta}{\sigma_{X_1}}(y_1-\overline{X}_1)\Phi(A_{y_1})-I\sigma_\theta\rho_m^2\rho_{XY}(1-\rho_m^2\rho_{XY}^2)^{-1/2}\varphi(A_{y_1})$$

and hence $\left.\dfrac{\partial Y_1^0}{\partial \rho_{XY}}=\dfrac{-\dfrac{\partial \Delta U(\rho_b,y_1,0)}{\partial \rho_{XY}}}{\dfrac{\partial \Delta U(\rho_b,y_1,0)}{\partial y_1}}\right]_{y_1=Y_1^0}=\dfrac{\rho_m\varphi(A_{Y_1})}{\dfrac{\rho_{XY}}{\sigma_{X_1}}(1-\rho_m^2\rho_{XY}^2)^{1/2}\Phi(A_{Y_1})}-2\dfrac{Y_1^0-\overline{X}_1}{\rho_{XY}}$, which is positive since

$Y_1^0 < X_1^0$ as indicated in the text and $X_1^0 < \overline{X}_1$ (see end of Appendix 3, where we introduced X_1^* and proved that

$X_1^F < X_1^* < \overline{X}_1$ for all values of F$\in[0,\mathrm{Ir}_2]$, in particular F=0). Finally, the increasingness of the cutoff in the

pattern of TRU and its decreasingness in the variance of the signal (i.e. σ_{X_1} for X_1^0, σ_{Y_1} for Y_1^0), coupled with

the fact that $\rho_b = \rho_m\rho_{XY} \le \rho_m$ and $\sigma_{Y_1}=\sigma_{X_1}/\rho_{XY}\ge\sigma_{X_1}$ ensures that $\lim_{\rho_{XY}\to 0}Y_1^0 < \lim_{\rho_m\to 0}X_1^0 < \lim_{\rho_m\to 0}X_1^* = -\infty$. It

should be obvious that $\lim_{\rho_{XY}\to 1}Y_1^0 = X_1^0$ for any ρ_m.

The next step is to prove Theorem 4. The new investment cutoff is $\max(X_1^F, Y_1^0)$ and the benefit from

contracting is $\max(Y_1^0 - X_1^F, 0)$. Using (A7) *mutatis mutandis*, we come to the conclusion that

$$\frac{\partial Y_1^0}{\partial \rho_m}=\frac{\partial Y_1^0}{\partial \rho_b}\frac{\partial \rho_b}{\partial \rho_m}=\frac{\sigma_{X_1}}{\left(1-\rho_m^2\rho_{XY}^2\right)^{1/2}}\cdot\frac{\varphi(A_{Y_1^0})}{\Phi(A_{Y_1^0})}-\frac{(Y_1^0-\overline{X}_1)}{\rho_m}>0 \text{ so that}$$

$$\frac{\partial(Y_1^0-X_1^F)}{\partial \rho_m}=\frac{\sigma_{X_1}}{\left(1-\rho_m^2\rho_{XY}^2\right)^{1/2}}\cdot\frac{\varphi(A_{Y_1^0})}{\Phi(A_{Y_1^0})}-\frac{\sigma_{X_1}}{\left(1-\rho_m^2\right)^{1/2}}\cdot\frac{\varphi(C_{X_1^F})}{\Phi(C_{X_1^F})}+\frac{(X_1^F-Y_1^0)}{\rho_m}$$

$$\le\frac{\sigma_{X_1}}{\left(1-\rho_m^2\right)^{1/2}}\left[\frac{\varphi(A_{Y_1^0})}{\Phi(A_{Y_1^0})}-\frac{\varphi(C_{X_1^F})}{\Phi(C_{X_1^F})}\right]\quad\text{when } Y_1^0\ge X_1^F.$$

Now, since the function $h: x\mapsto\varphi(x)/\Phi(x)$ is decreasing, it suffices to show that $A_{Y_1^0}>C_{X_1^F}$ to prove the

negativity of $\partial(Y_1^0-X_1^F)/\partial \rho_m$ when $Y_1^0\ge X_1^F$. This proof is omitted here but available upon request from the

author[26]. When $Y_1^0<X_1^F$, the reduction in risk-shifting is zero and therefore still non-increasing in ρ_m. □

Now, when ρ_{XY} tends to 1, Y_1^0 tends to X_1^0 and $\lim_{\rho_m\to 0}\left(Y_1^0-X_1^F\right)=\lim_{\rho_m\to 0}\left(X_1^0-X_1^F\right)=\infty$, while when ρ_{XY} is

arbitrarily small, $Y_1^0-X_1^F$ is arbitrarily negative for any value of ρ_m, in particular in the neighborhood of ρ_m=0.

---

[26] Attention: we stated earlier that $C=A-F/[I\sigma_\theta(1-\rho^2)^{1/2}]$. However, here $A_{Y_1^0}$ uses σ_Y and $\rho_b=\rho_m\rho_{XY}$, while $C_{X_1^F}$ uses

σ_X and ρ_m.

Since $\partial(Y_1^0 - X_1^F)/\partial\rho_{XY} > 0$ (Proposition 4), there is a critical $\underline{\rho_{XY}}$ such that $\lim_{\rho_m \to 0}\left(Y_1^0 - X_1^F\right) = 0$. Given that

$\partial(Y_1^0 - X_1^F)/\partial\rho_m < 0$ whenever $Y_1^0 \geq X_1^F$ (see previous paragraph), this ensures that for $\rho_{XY} \leq \underline{\rho_{XY}}$, Y_1^0 is

never greater than X_1^F *for any* ρ_m. In other words, if Y_1^0 is not greater than X_1^F at low ρ_m values, it will not either for larger values of ρ_m.

In the case where $\rho_{XY} > \underline{\rho_{XY}}$ (i.e. it is not the case that $Y_1^0 < X_1^F$ for all ρ_m), when both ρ_m and ρ_{XY} tend to 1,

$\lim_{\substack{\rho_m \to 1 \\ \rho_{XY} \to 1}}\left(Y_1^0 - X_1^F\right) = \lim_{\rho_m \to 1}\left(X_1^0 - X_1^F\right) = 0$ and the increasingness of Y_1^0 in ρ_{XY} ensures that there is no ρ_{XY} for which

Y_1^0 is consistently greater than X_1^F (i.e. for all ρ_m). Since $\lim_{\rho_m \to 0}\left(Y_1^0 - X_1^F\right) > 0$ (since $\rho_{XY} > \underline{\rho_{XY}}$),

$\lim_{\substack{\rho_m \to 1 \\ \rho_{XY} \to 1}}\left(Y_1^0 - X_1^F\right) = 0$, $\dfrac{\partial(Y_1^0 - X_1^F)}{\partial\rho_{XY}} > 0$ and $\partial(Y_1^0 - X_1^F)/\partial\rho_m < 0$, the Intermediate Value Theorem ensures that

there indeed exists a unique ρ_m^* such that, for a given ρ_{XY} greater than $\underline{\rho_{XY}}$, $Y_1^0 > X_1^F$ for $\rho_m < \rho_m^*$ and

$Y_1^0 \leq X_1^F$ for $\rho_m \geq \rho_m^*$. For $\rho_{XY} < \underline{\rho_{XY}}$, we'll set $\rho_m^* = 0$ and for $\rho_{XY} = 1$, $\rho_m^* = 1$. The improvement $Y_1^0 - X_1^F$ in

terms of risk-shifting is then decreasing in ρ_m on $[0, \rho_m^*)$ and constant (equal to 0) on $[\rho_m^*, 1]$.

It can also be shown that this ρ_m^* is increasing in ρ_{XY}: an implicit differentiation at $\rho_m = \rho_m^*$ yields, as $Y_1^0 - X_1^F$ is kept constant at 0:

$$\frac{\partial\rho_m^*}{\partial\rho_{XY}} = -\left.\frac{\partial Y_1^0/\partial\rho_{XY}}{\partial(Y_1^0 - X_1^F)/\partial\rho_m}\right]_{\rho_m = \rho_m^*} > 0$$

since the numerator is positive as stated in Proposition 4 and the denominator is negative as was just proven.

Alternatively, since $\rho_{XY} = \sigma_{X_1}/\sqrt{\sigma_{X_1}^2 + \sigma_\varepsilon^2}$, ρ_m^* is decreasing in σ_ε.

7. Partial Contracting, TRU and Corporate Debt Yields

♦ For a given ρ_{XY}, if $Y_1^0 \leq X_1^F$ (i.e. for all ρ_m if $\rho_{XY} \leq \underline{\rho_{XY}}$, for $\rho_m \geq \rho_m^*$ otherwise), nothing is changed and we refer the reader to Appendix 5 for a proof that bond yields are decreasing in ρ_m.

• if $\rho_{XY} > \underline{\rho_{XY}}$ and $\rho_m < \rho_m^*$, shareholders will invest, on average[27], for all $x_1 > Y_1^0$. This yields a bond price similar to the one in (A8) and a first derivative of bond prices with respect to ρ_m similar to (A9), but with Y_1^0 in lieu of the generic cutoff ξ. The reader should also be aware that in this derivative, $A_{Y_1^0}$ denotes

$$E(\theta \mid x_1 = Y_1^0)/\sigma_{\theta \mid X_1} = \left[\bar{\theta} + \rho_m \sigma_\theta (Y_1^0 - \bar{X}_1)/\sigma_{X_1}\right]/\left[\sigma_\theta(1-\rho_m^2)^{1/2}\right],$$ as opposed to Appendix 6, where it denoted

$$E(\theta \mid y_1 = Y_1^0)/\sigma_{\theta \mid Y_1} = \left[\bar{\theta} + \rho_b \sigma_\theta (Y_1^0 - \bar{X}_1)/\sigma_{Y_1}\right]/\left[\sigma_\theta(1-\rho_b^2)^{1/2}\right].$$ Since we consider here the case where

$Y_1^0 > X_1^F$, $\quad C_{Y_1^0}\Phi(C_{Y_1^0}) + \varphi(C_{Y_1^0}) > C_{X_1^F}\Phi(C_{X_1^F}) + \varphi(C_{X_1^F}) = (Ir_2 - F)/\left[I\sigma_\theta(1-\rho_m^2)^{1/2}\right]$ and since $Y_1^0 < X_1^0$,

$A_{Y_1^0}\Phi(A_{Y_1^0}) + \varphi(A_{Y_1^0}) < A_{X_1^0}\Phi(A_{X_1^0}) + \varphi(A_{X_1^0}) = r_2/\left[\sigma_\theta(1-\rho_m^2)^{1/2}\right]$; this, coupled with the fact that $\partial Y_1^0/\partial\rho_m > 0$ (see previous Appendix), is enough to show, through the exact same steps as in Appendix 5, that

$$\partial r_1 r_2 B_0(\rho_m, F, Y_1^0)/\partial\rho_m > I\sigma_\theta\left[\Phi(A_{Y_1^0}) - \Phi(C_{Y_1^0})\right]\varphi\left((Y_1^0 - \bar{X})/\sigma_{X_1}\right) > 0.$$ Hence bond prices are still increasing in ρ_m after optimal contracting; equivalently, the yield premium demanded is lower the quicker the uncertainty is resolved. $\square$

oooooo

It can also be shown that this yield premium is non-increasing in ρ_{XY}, the accuracy of the information Y_1 available to outsiders. As before, we will do that through looking at comparative statics involving bond prices:

• if $Y_1^0 \leq X_1^F$ (i.e. $\rho_m \geq \rho_m^*$), $\dfrac{\partial B_0(\rho_m, F, X_1^F)}{\partial\rho_{XY}} = 0$;

• if $Y_1^0 > X_1^F$ (i.e. $\rho_m < \rho_m^*$),

$$\frac{\partial r_1 r_2 B_0(\rho_m, F, Y_1^0)}{\partial\rho_{XY}} = \frac{F}{\sigma_{X_1}}\varphi\left(\frac{Y_1^0 - \bar{X}_1}{\sigma_{X_1}}\right) + I\sigma_\theta\left(1-\rho_m^2\right)^{1/2}\left[C^{Y_1^0}\Phi(C^{Y_1^0}) + \varphi(C^{Y_1^0}) - A^{Y_1^0}\Phi(A^{Y_1^0}) + \varphi(A^{Y_1^0})\right]\varphi_{X_1}(Y_1^0)\frac{\partial Y_1^0}{\partial\rho_{XY}}$$

Now, as noted before, $I\sigma_\theta(1-\rho_m^2)^{1/2}\left[C^{Y_1^0}\Phi(C^{Y_1^0}) + \varphi(C^{Y_1^0})\right] > Ir_2 - F$ since $Y_1^0 > X_1^F$ and

$I\sigma_\theta(1-\rho_m^2)^{1/2}\left[A^{Y_1^0}\Phi(A^{Y_1^0}) + \varphi(A^{Y_1^0})\right] < Ir_2$ since $Y_1^0 < X_1^0$; this, combined with the positivity of $\partial Y_1^0/\partial\rho_{XY}$,

yields the desired result: bond yields are decreasing in ρ_{XY} on $[0, \rho_m^*)$ and independent of ρ_{XY} on $[\rho_m^*, 1]$. $\square$

This effect is reinforced by the increasingness of ρ_m^* in ρ_{XY} (see previous Appendix): as the latter increases, the region where bond yields are strictly decreasing in ρ_{XY} expands.

---

[27] As we noted in the text, shareholders will really invest for all realizations of X_1 greater than $Y_1^0 - \varepsilon_1$. But since bondholders do not observe ε_1, they will price securities using $E(\varepsilon)=0$ for ε_1.

Finally, the price of the equity $S_0(\rho_m, F, \xi)$ can be expressed as the difference between the market value of the firm (see footnote 12), also increasing in ρ_{XY}, and the price of bonds (A8):

$$S_0(\rho_m, F, \xi) = \frac{1}{r_1 r_2}\left\{(Ir_2 - F)\Phi\left(\frac{\xi - \overline{X}}{\sigma_{X_1}}\right) + I\sigma_\theta(1-\rho_m^2)^{1/2}\int_\xi^\infty\left[C^{x_1}\Phi(C^{x_1}) + \varphi(C^{x_1})\right]dP(x_1)\right\}$$

where the cutoff ξ used is $\max(X_1^F, Y_1^0)$. Now, if $X_1^F > Y_1^0$, share prices are insensitive to a change in ρ_{XY} until ρ_{XY} is large enough for Y_1^0 to exceed X_1^F, in which case they become sensitive to a further increase in ρ_{XY}. Then

$$\frac{\partial r_1 r_2 S_0(\rho_m, F, Y_1^0)}{\partial \rho_{XY}} = \frac{1}{\sigma_{X_1}}\varphi\left(\frac{Y_1^0 - \overline{X}_1}{\sigma_{X_1}}\right)[Ir_2 - F]\frac{\partial Y_1^0}{\partial\rho_{XY}} - I\sigma_\theta\left(1-\rho_m^2\right)^{1/2}\left[C^{Y_1^0}\Phi(C^{Y_1^0}) + \varphi(C^{Y_1^0})\right]\varphi_{X_1}(Y_1^0)\frac{\partial Y_1^0}{\partial\rho_{XY}}$$

Given that $I\sigma_\theta(1-\rho_m^2)^{1/2}\left[C^{Y_1^0}\Phi(C^{Y_1^0}) + \varphi(C^{Y_1^0})\right] > Ir_2 - F$ since $Y_1^0 > X_1^F$, this last derivative is negative: as ρ_{XY} increases, contracting becomes more efficient, triggering a transfer of values from shares to bonds.

References

Akerlof, G. A., 1970, "The Market for 'Lemons': Qualitative Uncertainty and the Market Mechanism", *Quarterly Journal of Economics*, 84, 488-500.

Barclay, M. J., and C. W. Smith, 1995, "The Maturity Structure of Corporate Debt", *Journal of Finance*, 50, 609-631.

Bradley, M., G. A. Jarrell and H. Kim, 1984, "On the Existence of an Optimal Capital Structure: Theory and Evidence", *Journal of Finance*, 39, 857-880.

Brennan, M. J., 1995, "Corporate Finance Over the Past 25 Years", *Financial Management*, 24, 9-22.

Brito, J. A., and K. John, 2002, "Leverage and Growth Opportunities: Risk-Avoidance Induced by Risky Debt", Working Paper, Stern School of Business, New York University, January.

Bronte, S., 1982, *Japanese Finance: Markets and Institutions*, Euromoney Publications Limited, London.

Duffie, D., M. Schroder, and C. Skiadas, 1996, "Recursive Valuation of Defaultable Securities and the Timing of the Resolution of Uncertainty", *Annals of Applied Probability*, 6, 1075-1090.

Duffie, D. and D. Lando, 2001, "Term Structure of Credit Spreads with Incomplete Accounting Information", *Econometrica*, 69, 633-664.

Epstein, L. G., 1980, "Decision Making and the Temporal Resolution of Uncertainty", *International Economic Review*, 21, 269-283.

Epstein, L. G., and S. M. Turnbull, 1980, "Capital Asset Prices and the Temporal Resolution of Uncertainty", *Journal of Finance*, 35, 627-643.

Goswami, G., T. Noe, and M. Rebello, 1995, "Debt financing under Asymmetric Information", *Journal of Finance*, 50, 633-659.

Guedes, J., and T. Opler, 1996, "The Determinants of the Maturity of Corporate Debt Issues", *Journal of Finance*, 51, 1809-1833.

Jensen, M. C., and W. H. Meckling, 1976, "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure", *Journal of Financial Economics*, 3, 305-360.

John, K., 1987, "Risk-Shifting Incentives and Signaling Through Corporate Capital Structure", *Journal of Finance*, 42, 623-641.

John, K., and J. Ronen, 1990, "Information Structures, Optimal Contracts and the Theory of the Firm", *Journal of Accounting, Auditing and Finance*, 5, 61-95.

Kester, C. W., 1986, "Capital and Ownership Structure: A Comparison of United States and Japanese Manufacturing Corporations", *Financial Management*, 5-16.

Leland, H. E., 1994, "Corporate Debt Value, Bond Covenants and Optimal Capital Structure", *Journal of Finance*, 49, 1213-1252.

Leland, H. E., 1998, "Agency Costs, Risk Management and Capital Structure", *Journal of Finance*, 53, 1213-1243.

Long, M. S., and I. B. Malitz, 1986, "Investment Patterns and Financial Leverage", NBER Study edited by Benjamin Friedman, University of Chicago Press, 325-348.

Modigliani, F., and M. H. Miller, 1958, "The Cost of Capital, Corporation Finance and the Theory of Investment", *American Economic Review*, 48, 261-297.

Myers, S. C., 1977, "Determinants of Corporate Borrowing", *Journal of Financial Economics*, 5, 147-175.

Nabar, P., R. C. Stapleton, and M. G. Subrahmanyam, 1988, "Default Risk, Resolution of Uncertainty and the Interest Rate on Corporate Loans", *Studies in Banking and Finance*, 5, 221-245.

Reisz, A. and C. Perlich, 1999, "Temporal Resolution of Uncertainty and Corporate Debt Yields: an Empirical Investigation", Working Paper FIN-99-043, Stern School of Business, New York University.

Ross, S., 1978, "A Simple Approach to Valuation of Risky Streams", *Journal of Business*, 51, 453-475.

Ross, S., 1989, "Information and Volatility: the No-arbitrage Martingale Approach to Timing and Resolution Irrelevancy", *Journal of Finance*, 44, 1-17.

Rothschild, M., and J. Stiglitz, 1970, "Increasing Risk I: A Definition", *Journal of Economic Theory*, 2, 225-243.

Rothschild, M., and J. Stiglitz, 1971, "Increasing Risk II: Its Economic Consequences", *Journal of Economic Theory*, 3, 66-84.

Rubinstein, M., and P. Carr, 1995, "Appendix: Useful Facts about Normal Distributions", in Rubinstein, M., and E. Reiner, "Exotic Options", Working Paper, Haas School of Business, University of California at Berkeley.

Smith, C. W. Jr., and R. Stulz, 1985, "The Determinants of Firm's Hedging Policies", *Journal of Financial and Quantitative Analysis*, 20, 391-405.

Stapleton, R. C., and M. G. Subrahmanyam, 1984, "The Valuation of Multivariate Contingent Claims in Discrete Time Models", *Journal of Finance*, 39, 3, 207-228.

Stohs, M. H., and D. C. Mauer, 1996, "The Determinants of Corporate Debt Maturity Structure", *Journal of Business*, 69, 279-311.

Titman, S., 1984, "The Effect of Capital Structure on a Firm's Liquidation Decision", *Journal of Financial Economics*, 13, 137-151.

Titman, S., and R. Wessels, 1988, "The Determinants of Capital Structure Choice", *Journal of Finance*, 43, 1-19.
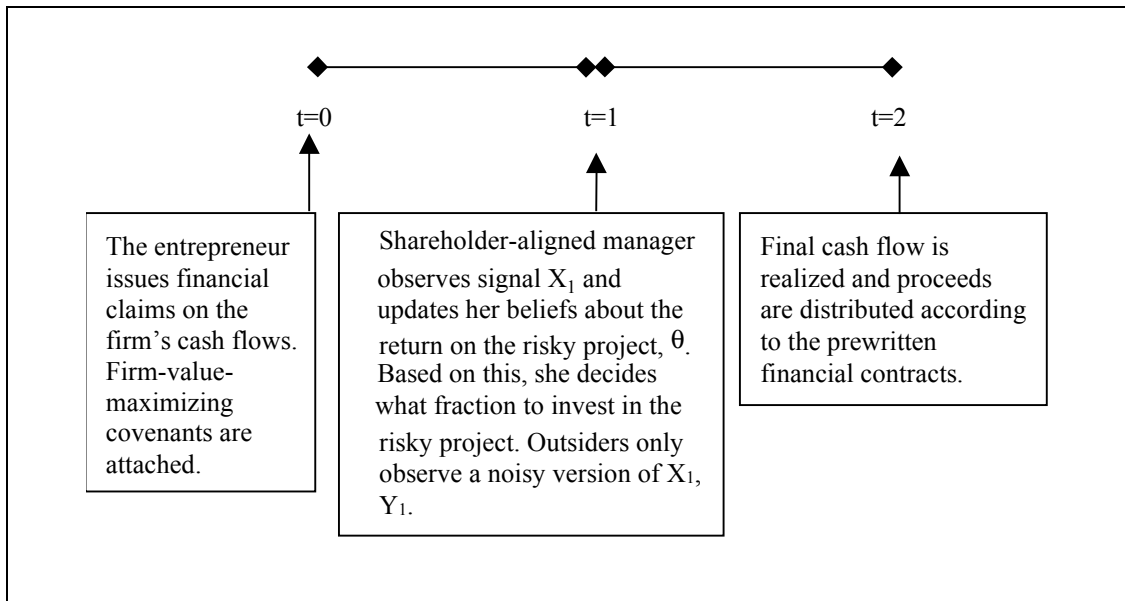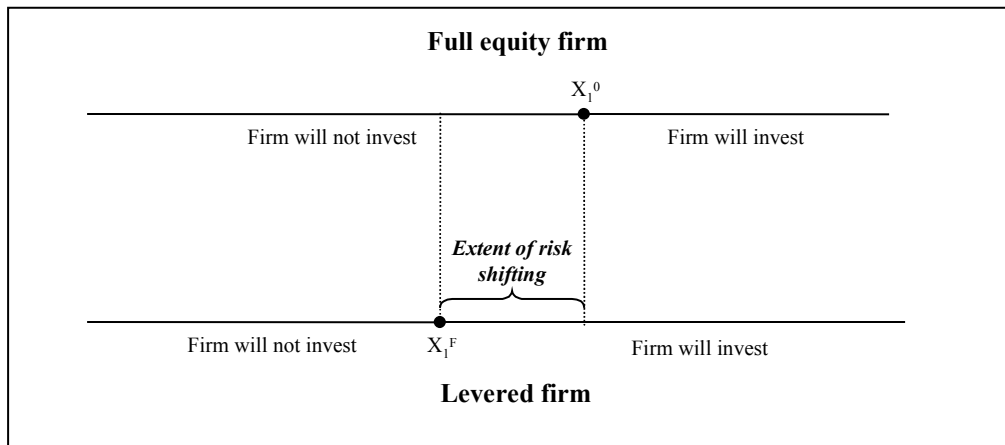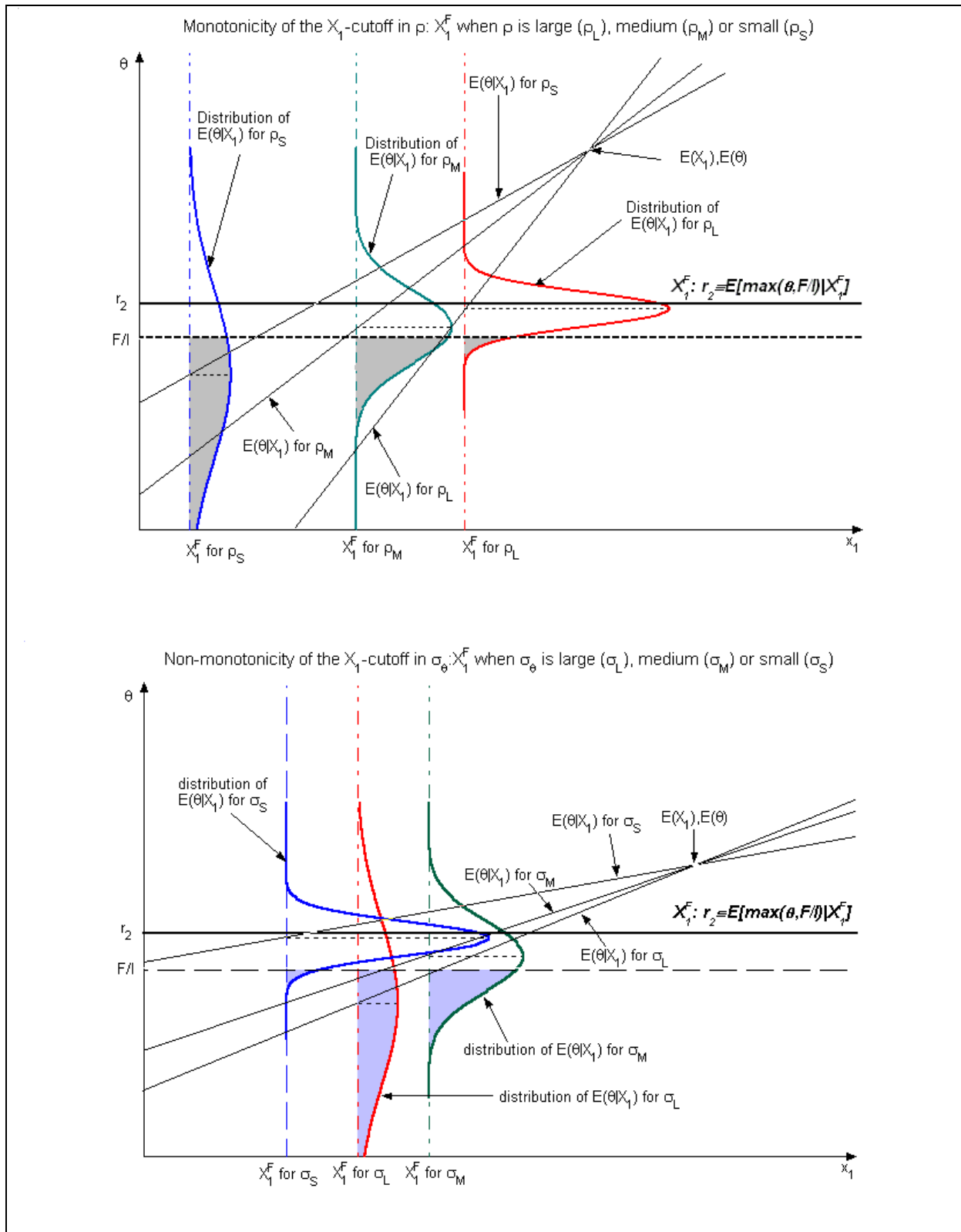
Figures

The entrepreneur issues financial claims on the firm's cash flows. Firm-value-maximizing covenants are attached.

Shareholder-aligned manager observes signal X_1 and updates her beliefs about the return on the risky project, θ. Based on this, she decides what fraction to invest in the risky project. Outsiders only observe a noisy version of X_1, Y_1.

Final cash flow is realized and proceeds are distributed according to the prewritten financial contracts.

t=0   t=1   t=2

Figure 1

Full equity firm

X_1^0

Firm will not invest   Firm will invest

Extent of risk shifting

Firm will not invest X_1^F Firm will invest

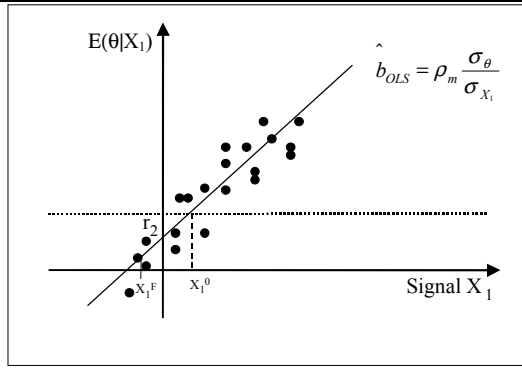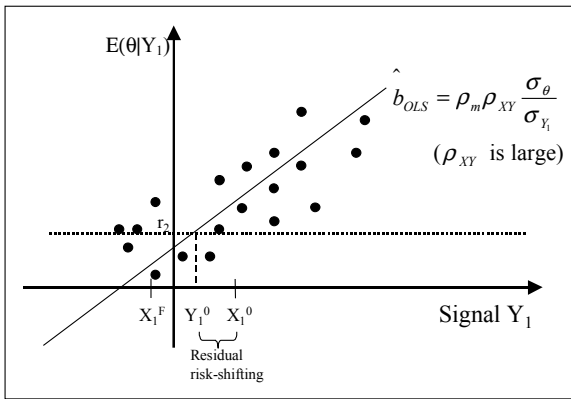**Levered firm**

Figure 2

Figure 3
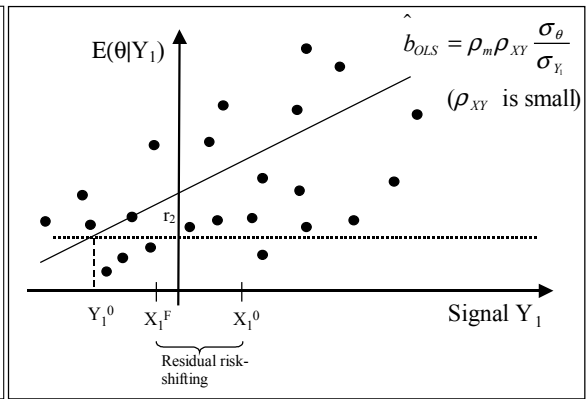
cutoff of an all-equity firm

contracting reduces the extent of risk-shifting   contracting doesn't help

Figure 4

Figure 5
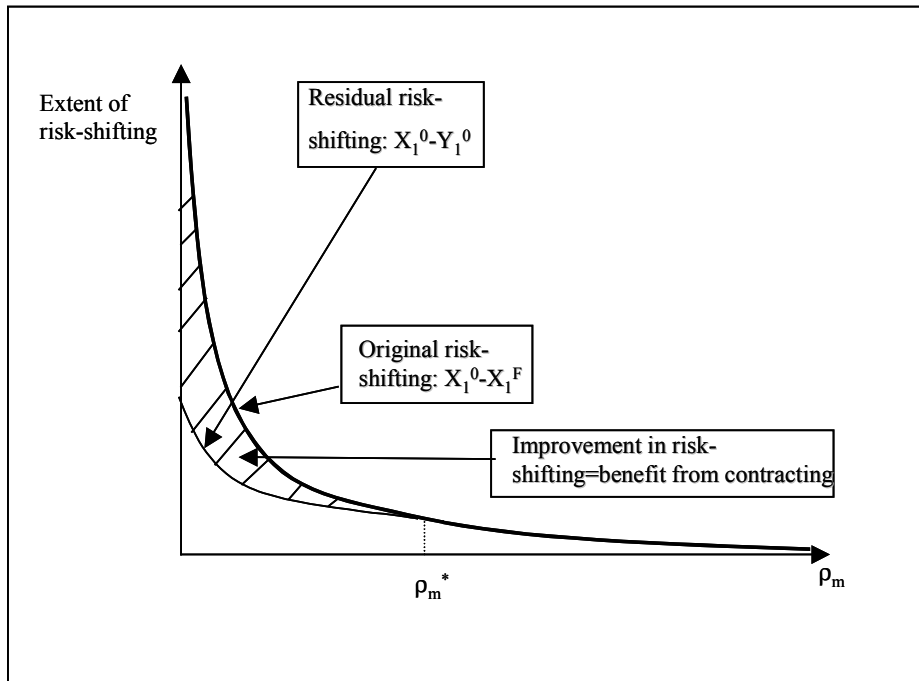
Agency costs as a function of the pattern of TRU and face value of debt when σ_θ=0.6, r_2=1.05, and $E(\theta)$=1.15

Corporate debt yield premium as a function of the pattern of TRU and face value of debt when σ_θ=0.6, r_2=1.05, and $E(\theta)$=1.15

Residual risk-shifting after optimal contracts have been written when F=50, r_2=1.05, $E(\theta)$=1.15, and σ_θ=0.6

Residual agency costs after optimal contracting when σ_θ=0.6, r_2=1.05, $E(\theta)$=1.15, and F=50

The maximum ρ_m for which contracting still presents benefits when F=50, r_2=1.05, and $E(\theta)$=1.15

The maximum ρ_m for which contracting still presents benefits when σ_θ=0.6, r_2=1.05, and $E(\theta)$=1.15